Session 3: The rank

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Contents

The rank

Recap of the videos

Definition

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

$$\mathrm{rank}(x_1, \dots, x_k) \stackrel{\mathrm{def}}{=} \dim(\mathrm{Span}(x_1, \dots, x_k)).$$

Definition

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

$$\mathrm{rank}(M) \stackrel{\mathrm{def}}{=} \mathrm{rank}(c_1, \dots, c_m) = \dim(\mathrm{Im}(M)).$$

Proposition

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

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

How do we compute the rank ?

$\alpha \neq 0$

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

$$\operatorname{rank}(v_1, \ldots, v_k) = \begin{cases} \operatorname{rank}(v_1, \ldots, v_{i-1}, \alpha v_i, v_{i+1}, \ldots, v_k) \\ \rightarrow \text{multiply one vector by } \alpha \neq 0 \\ \operatorname{rank}(v_1, \ldots, v_{i-1}, \underbrace{v_i + \beta v_j}, v_{i+1}, \ldots, v_k) \end{cases}$$

$\rightarrow$ replace v_i by $v_i + \beta v_j$ for some $j \neq i$

As a consequence, the Gaussian elimination method keeps the rank of a matrix unchanged!

Example

Let's compute the rank of $A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ -1 & 5 & 2 & 0 \end{pmatrix}$ $\begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix}$

$A' = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & -3 \\ 0 & 4 & 2 & 1 \end{pmatrix}$ $\begin{matrix} \\ R_2 - 2R_1 \\ R_3 + R_1 \end{matrix}$

here: $\operatorname{rank}(A) = \operatorname{rank}(A')$

$A'' = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & -3 \\ 0 & 0 & 0 & 7 \end{pmatrix}$ $\begin{matrix} \\ \\ R_3 - 2R_2 \end{matrix}$

$\begin{matrix} 0 & 0 & 0 & 1 \end{matrix}$

$\operatorname{rank}(A'') = \operatorname{rank}(A)$

but

$\dim \operatorname{Im} A'' = \dim \operatorname{Im} A$

Example

Claim: $\mathrm{rank}(A'') = 3$

$$A'' = \begin{array}{cccc} c_1 & c_2 & & c_4 \\ \end{array}$$
$$\begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & -3 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

"leading coefficients"

Claim: c_1, c_2, c_4 are linearly independent

Let $a, b, c \in \mathbb{R}$ such that $\underbrace{a c_1 + b c_2 + c c_4 = 0}$

$$\begin{cases} a - b + c = 0 \\ 2b - 3c = 0 \\ 7c = 0 \end{cases} \iff \begin{cases} a = 0 \\ b = 0 \\ c = 0 \end{cases}$$

Hence c_1, c_2, c_4 are lin indep.

$\mathrm{Span}(c_1, c_2, c_4) = \mathbb{R}^3$ & $\mathrm{Im}(A'') = \mathbb{R}^3$

The rank-nullity Theorem

Rank-nullity Theorem

Theorem

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

$$\underbrace{\operatorname{rank}(L)}_{\dim(\operatorname{Im}(L))} + \dim(\operatorname{Ker}(L)) = \underbrace{m}_{\dim(\mathbb{R}^m)}.$$

Very usefull to get dim Ker L from rank(L) and vice-versa !

Intuition

Let us solve the linear system $Ax = 0$.

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 0 \\ 2 & 0 & 1 & -1 & 0 \\ -1 & 5 & 2 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 0 \\ 0 & 2 & 1 & -3 & 0 \\ 0 & 4 & 2 & 1 & 0 \end{array}\right) \begin{array}{l} (R_1) \\ (R_2) - 2(R_1) \\ (R_3) + (R_1) \end{array}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 0 \\ 0 & 2 & 1 & -3 & 0 \\ 0 & 0 & 0 & 7 & 0 \end{array}\right) \begin{array}{l} (R_1) \\ (R_2) \\ (R_3) - 2(R_2) \end{array} \to \begin{cases} x_1 - x_2 + x_4 = 0 \\ 2x_2 + x_3 - 3x_4 = 0 \\ 7x_4 = 0 \end{cases}$$

$$\begin{cases} x_1 - x_2 = 0 \\ 2x_2 = -x_3 \\ x_4 = 0 \end{cases} \qquad \begin{cases} x_1 = -\frac{x_3}{2} \\ x_2 = -x_3/2 \\ x_3 = x_3 \\ x_4 = 0 \end{cases}$$

$$\left\{ \begin{pmatrix} -t/2 \\ -t/2 \\ t \\ 0 \end{pmatrix} \middle| t \in \mathbb{R} \right\} = \mathrm{Span}\begin{pmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix} = \mathrm{Ker}(A)$$

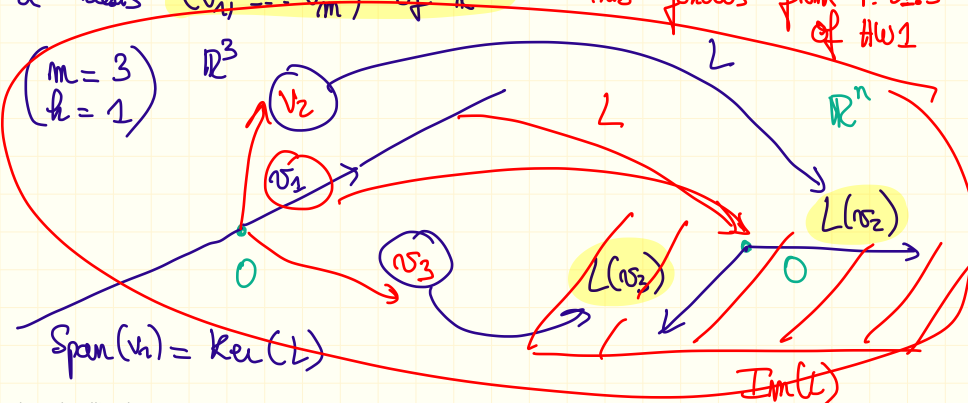
$\dim \mathrm{Ker}(A) = 4 -$ "number of leading coefs"
$= 4 - \mathrm{rank}(A)$

Proof of the rank-nullity Theorem

- Let $k = \dim \mathrm{Ker}(L)$ and $(v_1, \dots v_k)$ be a basis of $\mathrm{Ker}(L)$.

- I can add vectors $v_{k+1}, \dots v_m$ to it to obtain a basis $(v_1, \dots v_m)$ of $\mathbb{R}^m$

this follows from P. 1.3 of HW1



Proof of the rank-nullity Theorem

Claim: $(L(v_{k+1}) \ldots L(v_m))$ is a basis of $\mathrm{Im}(L)$

Given the claim, the theorem follows.

$$\dim \mathrm{Im}(L) = m - k = m - \dim \mathrm{Ker}(L)$$

① $\underbrace{\mathrm{Span}(L(v_{k+1}) \ldots L(v_m))}_{S} = \mathrm{Im}(L)$.

- $S \subset \mathrm{Im}(L)$ because $L(v_{k+1}) \ldots L(v_m) \in \mathrm{Im}(L)$ and $\mathrm{Im}(L)$ is a subspace.

Proof of the rank-nullity Theorem

- $\mathrm{Im}(L) \subset \underbrace{\mathrm{Span}(L(v_{k+1}) \_\_ L(v_m))}_{S}$
- Let $y \in \mathrm{Im}(L)$, there exists $x \in \mathbb{R}^m$ such that $y = L(x)$.
- let $(\alpha_1 \_\_ \alpha_m)$ be the coord. of x in $(v_1 \_\_ v_m)$

$$y = L(x) = L(\alpha_1 v_1 + \dots + \alpha_m v_m)$$

$$= \alpha_1 \underbrace{L(v_1)}_{=0} + \underbrace{\dots +}_{=0} \alpha_k \underbrace{L(v_k)}_{=0} + \underbrace{\alpha_{k+1} L(v_{k+1}) + \dots + \alpha_m L(v_m)}_{\in S}$$

$\boxed{y \in S}$

hence $\mathrm{Im}(L) = \mathrm{Span}(L(v_{k+1}) \dots, L(v_m))$

Proof of the rank-nullity Theorem

② Let's show that $\boxed{L(v_{k+1}) \_\_ L(v_m) \text{ lin. indep.}}$

Let $\alpha_{k+1} \dots \alpha_m \in \mathbb{R}$ such that $\alpha_{k+1} L(v_{k+1}) + \dots + \alpha_m L(v_m) = 0$

$$\underline{0} = L(\underbrace{\alpha_{k+1} v_{k+1} + \dots + \alpha_m v_m}_{u}) = \alpha_1 v_1 + \dots + \alpha_k v_k$$

$\leftarrow L(u) = 0$

Hence $u \in \operatorname{Ker}(L)$ therefore there exists $\alpha_1 \dots \alpha_k$ s.t.

$$u = \alpha_1 v_1 + \dots + \alpha_k v_k$$

$-u$

we get: $\alpha_1 v_1 + \dots + \alpha_k v_k - \alpha_{k+1} v_{k+1} - \dots - \alpha_m v_m = 0$

$(v_1 \_\_ v_m)$ lin. indep hence $\alpha_1 = \dots = \alpha_m = 0$.

Proof of the rank-nullity Theorem

Invertible matrices

Invertible matrices

Definition (Matrix inverse)

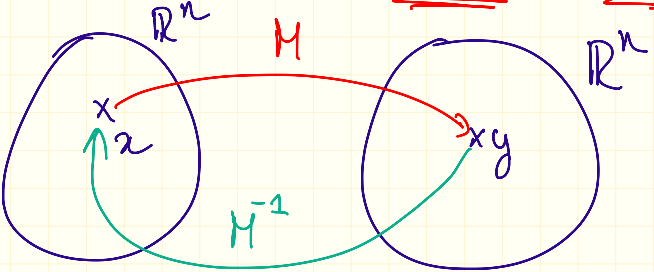
A **square** matrix $M \in \mathbb{R}^{n\times n}$ is called *invertible* if there exists a matrix $M^{-1} \in \mathbb{R}^{n\times n}$ such that

$$MM^{-1} = M^{-1}M = \mathrm{Id}_n.$$

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Such matrix M^{-1} is unique and is called the *inverse* of M.

Exercise: Let $A, B \in \mathbb{R}^{n\times n}$. Show that if $AB = \mathrm{Id}_n$ then $BA = \mathrm{Id}_n$.



Invertible matrices

Theorem

Let $M \in \mathbb{R}^{n \times n}$. The following points are equivalent:

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

- The rank nullity theorem gives (2) $\iff$ (3)
- (1) $\implies$ $\begin{pmatrix} (2) \\ (3) \end{pmatrix}$ $\implies$ (4) $\implies$ (1)

Proof

(1) $\Rightarrow$ (3) Assume M invertible.

Let $x \in \operatorname{Ker} M$: $Mx = 0$

$$\rightarrow \quad \underbrace{M^{-1}M}_{Id_n} x = M^{-1}0 = 0$$

$$x = Id_n\, x = 0 \qquad \boxed{x = 0}$$

this gives $\boxed{\operatorname{Ker}(L) = \{0\}}$

Proof

$\begin{pmatrix}(2)\\(3)\end{pmatrix} \Rightarrow (4)$ Let's assume that $\left| \begin{array}{l} \text{rank}(M) = n \\ \underline{\text{Ker}(M) = \{0\}} \end{array} \right.$

$\text{Im}(M)$ is a subspace of $\mathbb{R}^n$ $\quad \dim(\mathbb{R}^n) = n$

$\dim \text{Im}(M) = n$

Hence $\boxed{\text{Im}(M) = \mathbb{R}^n}$

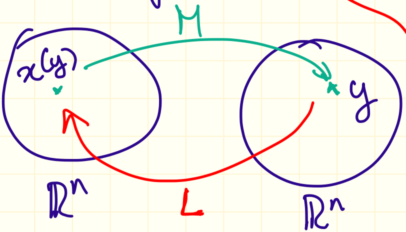
- From what we have seen last week, for every $y \in \mathbb{R}^n$, there exists (because $\text{Im}(M) = \mathbb{R}^n$) a unique (because $\text{Ker}\, M = \{0\}$) such that $Mx = y$

Proof

$(4) \Rightarrow (1)$ For every $y \in \mathbb{R}^n$ there exists $x(y) \in \mathbb{R}^n$ such that $M\,x(y) = y$.

Let's define:

$$L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
$$y \longmapsto x(y)$$



Claim: L is linear.

Proof

By construction of L, we have for all $y \in \mathbb{R}^n$

$$M\,L(y) = y$$

$$Id_n : \mathbb{R}^n \to \mathbb{R}^n$$
$$y \mapsto y$$

$$M \circ L\,(y) = Id_n\,(y)$$

This gives that the linear transformations $M \circ L$ and Id_n are the same : $M \circ L = Id_n$

Hence their matrices are equal:

"a matrix product"

$$\boxed{\tilde{M}\,\tilde{L} = \tilde{I}d_n}$$

: $\tilde{M}$ is invertible

Transpose of a matrix

Transpose of a matrix

Definition

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

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

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

Example: $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ $\quad M^\mathsf{T} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$

"the col. of M become the rows of M^T"

Remark:

- We have $(M^\mathsf{T})^\mathsf{T} = M$.
- The mapping $M \mapsto M^\mathsf{T}$ is linear.

$(A+B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T}$

$(\alpha A)^\mathsf{T} = \alpha A^\mathsf{T}$

Properties of the transpose

Proposition

For all $A \in \mathbb{R}^{n\times m}$, $\boxed{\operatorname{rank}(A) = \operatorname{rank}(A^\mathsf{T}).}$

because "rank of rows" = "rank of cols."

proved in recitation

Proposition

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

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

Proof. Let's compute

$$\left((AB)^T\right)_{i,j} = (AB)_{j,i} = \sum_{\ell=1}^{m} A_{j,\ell} B_{\ell,i}$$

$$= \sum_{\ell=1}^{m} (A^T)_{\ell,j} (B^T)_{i,\ell} = \sum_{\ell=1}^{m} (B^T)_{i,\ell} (A^T)_{\ell,j}$$

$$= (B^T A^T)_{i,j}$$

□

Symmetric matrices

Definition

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

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

or, equivalently if $A = A^\mathsf{T}$.

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

$(MM^T)^T = (M^T)^T M^T = MM^T \quad \therefore MM^T$ is sym.

Ex: $\begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{pmatrix}$ is symmetric

Is the rank useful in practice?

Back to the movies ratings example

Assume that you are given the matrix of movies ratings:

users

$$\begin{pmatrix} 1 & 1 & 5 & 5 & 5 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1.001 & 5 & 5 & 5 \\ 2 & 2 & 2 & 0.0001 & 0 \\ 2.0001 & 2 & 2 & 0 & 0 \end{pmatrix}$$

M

Goal: how many different « user profiles » do we have ?

$\mathrm{rank}(M) = 5$

Conclusion

- The rank is not «robust» !
- We need to have a way to check if a matrix has «approximately a small rank».
- Equivalentely, given m vectors, one would like to be able to see if there exists a subspace of dimension $k \ll m$ from which the vectors are « close ».
- The singular value decomposition (lecture 6-7) will solves our problems !

Questions?

$$U \cap V = \{0\}$$

$(u_1 - u_k)$ $(v_1 - v_m)$

$(u_1 \ldots u_k \; v_1 \ldots v_m)$ lin indep.

$$\alpha_1 u_1 + \ldots + \alpha_k u_k + \beta_1 v_1 + \ldots + \beta_m v_m = 0$$

$\rightarrow \alpha_i = \beta_i$ are all zero

Questions?

MM^T is always symmetric

Let's apply this to $M = A^T$

$\underbrace{A^T (A^T)^T}_{A^T A}$ is ~~also~~ symmetric

$$\begin{array}{cc} & 2 \\ 10 & \begin{pmatrix} \\ M \\ \\ \end{pmatrix} \end{array} \qquad \begin{array}{cc} & 10 \\ 2 & \begin{pmatrix} M^T \end{pmatrix} \end{array}$$

$$MM^T = \begin{array}{cc} 10 & \\ \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} & 10 \end{array}$$

$$M^T M = \begin{array}{cc} 2 & \\ \begin{pmatrix} \; \end{pmatrix} & 2 \end{array}$$

Questions?

MM^T is 10×10 when M is 10×2.

MM^T is not invertible!

$\text{rank}(MM^T) \leq \min(\text{rank}\, M, \text{rank}\, M^T) \leq 2$

$$\iff \begin{cases} \text{rank}(MM^T) \leq \text{rank}(M) \\ \text{rank}(MM^T) \leq \text{rank}(M^T) \end{cases}$$