Session 3: Matrix Rank

Optimization and Computational Linear Algebra for Data Science

Marylou Gabrié (based on material by Léo Miolane)

Contents

- 1. The rank
 - 1.1 Definition
 - 1.2 Intuition from Data Science
 - 1.3 How do we compute the rank?
- 2. The rank-nullity theorem
 - 2.1 Theorem
 - 2.2 Inequalities
- 3. Rank and invertible matrices
- Transpose
 - 4.1 Definition and basic properties
 - 4.2 Symmetric Matrices
- 5. Why do we care?

Low rank matrices

1. The rank

1. The rank 3/27

Rank of a family of vectors

Definition

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

$$\operatorname{rank}(x_1,\ldots,x_k) \stackrel{\text{def}}{=} \dim(\operatorname{Span}(x_1,\ldots,x_k)).$$

1. The rank 1.1 Definition 4/27

1.1 Rank of a matrix: Definition

Definition

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

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

1. The rank 1.1 Definition 5/2

Example

1. The rank 1.1 Definition 6/27

« Rank of columns = rank of rows »

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

$$rank(r_1,\ldots,r_n)=rank(c_1,\ldots,c_m)=rank(M).$$

1. The rank 1.1 Definition 7/2

1.2 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 $\operatorname{«}\operatorname{rank}(M) = 5$ »?

1.2 Intuition from Data Science

Imagine now that

- ightharpoonup The rows of M corresponds to Netflix's users.
- ightharpoonup The columns of M corresponds to Netflix's movies.
- The entry $M_{i,j}$ is rating of the movie j by the user i, assuming that all the users have rated all the movies.

1.2 Intuition from Data Science

Imagine now that

- ightharpoonup The rows of M corresponds to Netflix's users.
- ightharpoonup The columns of M corresponds to Netflix's movies.
- The entry $M_{i,j}$ is rating of the movie j by the user i, assuming that all the users have rated all the movies.

Claim: the rank of M is "small".

- The ratings of a user can be obtained as a linear combination of a small number of « profiles ».
- In practice, we do not have access to the full matrix, so we can use this assumption to predict the missing entries.

1.3 How do we compute the rank?

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

$$rank(v_1, ..., v_k) = \begin{cases} rank(v_1, ..., v_{i-1}, \alpha v_i, v_{i+1}, ..., v_k) \\ \\ rank(v_1, ..., v_{i-1}, v_i + \beta v_j, v_{i+1}, ..., v_k) \end{cases}$$

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

Example

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

1. The rank 1.3 How do we compute the rank?

Example

1. The rank 1.3 How do we compute the rank?

12/27

2. The rank-nullity theorem

2.1 The rank-nullity theorem

Theorem

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

$$rank(L) + \dim(Ker(L)) = m.$$

Intuition

Let us solve the linear system Ax = 0 characterizing $x \in \ker(A)$.

$$\begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ -1 & 5 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Pro	of of	the ra	nk-n	ullity	y The	orem	
Eve	rcise.						
LXC	il Cise.						

2. The rank-nullity theorem 2.1 Theorem

16/27

2.2 Inequalities

Proposition

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

- 1. $\operatorname{rank}(A) \leq \min(n, m)$.
- 2. $rank(AB) \le min(rank(A), rank(B))$.

Proof.

2.2 Inequalities

Proposition

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

- 1. $\operatorname{rank}(A) \leq \min(n, m)$.
- 2. $rank(AB) \le min(rank(A), rank(B))$.

Proof.

Rank of invertible matrices

Theorem

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

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

Pi	ro	0	f											

Pi	ro	0	f											

Pi	ro	0	f											

Pi	ro	0	f											

4. Transpose of a matrix

4. Transpose of a matrix

4.1 Transpose of a matrix: Definition

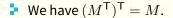
Definition

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 $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.

Remark:



The mapping $M \mapsto M^{\mathsf{T}}$ is linear.

Properties of the transpose

Proposition

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

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.

4.2 Symmetric matrices

Definition

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

5. Is the rank useful in practice?

Back to the movies ratings example

Assume that you are given the matrix of movies ratings:

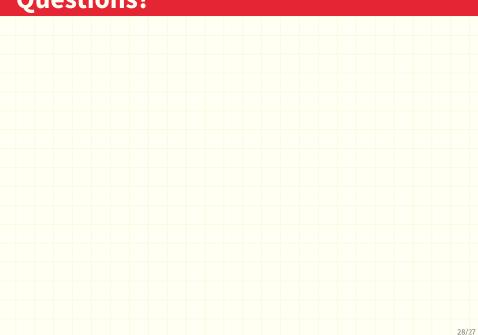
$$\begin{pmatrix} 1 & 1 & 5 & 5 & 5 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1.001 & 5 & 5 & 5 \\ 2 & 2 & 2 & 0.0001 & 0 \\ 2.0001 & 2 & 2 & 0 & 0 \end{pmatrix}$$

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

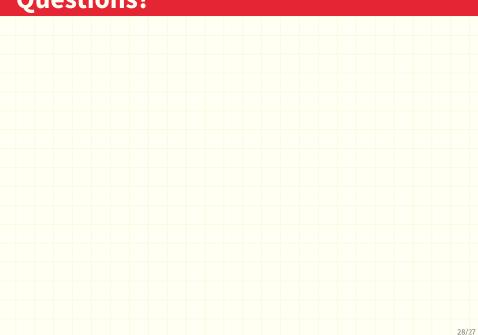
Conclusion

- 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!

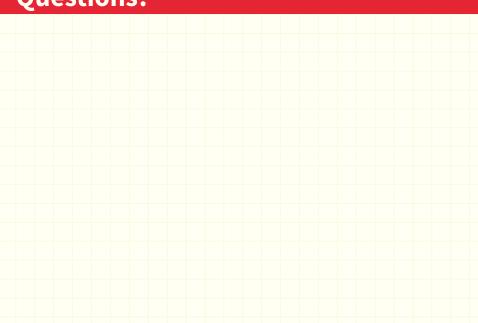
Questions?



Questions?



Questions?



28/27