AI511/MM505 Linear Algebra with Applications

Lecture 11 – Rank-Nullity Theorem, Eigenvalues and Eigenvectors

Vaidotas Characiejus

October 20, 2025

Department of Mathematics and Computer Science
University of Southern Denmark

Outline

Recap

Definition 3.1.2 of Johnston (2021)

Let $\mathcal{S}$ be a subspace of $\mathbb{R}^n$ with bases $B = \{v_1, v_2, \ldots, v_k\}$ and C. The *change-of-basis matrix* from B to C, denoted by $P_{C \leftarrow B}$,is the $k \times k$ matrix whose columns are the coordinate vectors $[v_1]_C$, $[v_2]_C$, . . ., $[v_k]_C$:

$$P_{C \leftarrow B} := \begin{bmatrix} [v_1]_C & [v_2]_C & \cdots & [v_k]_C \end{bmatrix}.$$

Theorem 3.1.3 of Johnston (2021)

Suppose B and C are bases of a subspace $\mathcal{S}$ of $\mathbb{R}^n$, and let $P_{C \leftarrow B}$ be the change-of-basis matrix from B to C. Then

(a) $P_{C \leftarrow B}[v]_B = [v]_C$ for all $v \in \mathcal{S}$, and

(b) $P_{C \leftarrow B}$ is invertible and $P_{C \leftarrow B}^{-1} = P_{B \leftarrow C}$.

Furthermore, $P_{C \leftarrow B}$ is the unique matrix with property (a).

Rank and nullity

Definition 2.4.4 of Johnston (2021)

The *rank* of $A \in \mathbb{R}^{m \times n}$ is denoted by $\mathrm{rank}(A)$ and defined by

$$\mathrm{rank}(A) = \dim(\mathrm{range}(A)).$$

The rank of a matrix can be thought of as a measure of how non-invertible a square matrix is (i.e., how small the subspace to which it squashes $\mathbb{R}^n$ is).

Definition

The *nullity* of $A \in \mathbb{R}^{m \times n}$ is denoted by $\mathrm{nullity}(A)$ and defined by

$$\mathrm{nullity}(A) = \dim(\mathrm{null}(A)).$$

Invertibility of a matrix

Theorem 2.4.9 of Johnston (2021)

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

(a) A is invertible;

(b) $\operatorname{rank}(A) = n$;

(c) $\operatorname{nullity}(A) = 0$.

Rank-nullity theorem

Relationship between $\mathsf{rank}(A)$ and $\mathsf{nullity}(A)$

- Recall that multiplying A by the vector x is equivalent to forming a linear combination of the columns of A, using the entries of x as the weights.
- If there are more ways to linearly combine the columns of A into the 0 vector, $\mathsf{null}(A)$ and $\mathsf{nullity}(A)$ are larger.
- Intuitively, the more ways you can combine the columns of A to get zero, the fewer ways there should be to combine them to reach any other non-zero vector.
- There should be a relationship between $\mathsf{rank}(A)$ and $\mathsf{nullity}(A)$. How does it look like?

Rank–nullity theorem

Theorem 2.4.10 of Johnston (2021)

Suppose that $A \in \mathbb{R}^{m\times n}$. Then

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = n.$$

Rank–nullity theorem (cont.)

$$\mathrm{rank}(A) = 3$$

$$\mathrm{rank}(A) = 3 \qquad \begin{bmatrix} 0 & 1 & 0 & * & * & 0 & * & * & * \\ 0 & 0 & 1 & * & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathrm{nullity}(A) = 6$$

- The rank of a matrix equals the number of leading columns that its reduced REF has, which equals the number of non-zero rows.
- Its nullity is the number of non-leading columns in its reduced REF, which equals the number of free variables.

Figure 2.24 of Johnston (2021)

Rank inequality and full rank

- We have that
$$\mathrm{rank}(A) \leq \min\{m, n\}$$
for $A \in \mathbb{R}^{m \times n}$ (think why this inequality is true).
- We say that A has *full rank* if $\mathrm{rank}(A) = \min\{m, n\}$.

Rank and matrix addition and multiplication

Theorem 2.4.11 of Johnston (2021)

Let A and B be matrices (with sizes such that the operations below make sense). Then

(i) $\operatorname{rank}(A+B) \le \operatorname{rank}(A) + \operatorname{rank}(B)$;

(ii) $\operatorname{rank}(AB) \le \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$.

Similar matrices

Similar matrices

Definition 3.1.3 of Johnston (2021)

We say that two matrices $A, D \in \mathbb{R}^{n \times n}$ are similar if there exists an invertible $P \in \mathbb{R}^{n \times n}$ such that $A = PDP^{-1}$.

- D is an arbitrary matrix in the definition (not necessarily diagonal).
- For example,

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & -1 \\ -2 & 0 \end{bmatrix}$$

are similar with

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

What does it mean that two matrices are similar?

- Let $p_1, \ldots, p_n \in \mathbb{R}^n$ be the columns of the invertible matrix P and

$$B = \{p_1, \ldots, p_n\}$$

  be a basis of $\mathbb{R}^n$ (think why this is the case).
- It follows that

$$P = P_{C \leftarrow B} \quad \text{and} \quad P^{-1} = P_{B \leftarrow C},$$

  where C is the standard basis of $\mathbb{R}^n$.
- The expression

$$A = PDP^{-1}$$

  shows that the transformation A in the standard basis is the same as the transformation D in the basis B.
- So what? The transformation D might be much simpler (diagonal) than the transformation A.

Motivating example

- Suppose we want to compute A^k, where $A \in \mathbb{R}^{n\times n}$ and k is a very large integer.
- If A is a general matrix (i.e., no special structure), computing A^k requires a lot of matrix multiplications and hence is computationally intensive.
- If A is a diagonal matrix, computing A^k is fairly straightforward since

$$A^k = \mathrm{diag}(a_{11}^k, a_{22}^k, \ldots, a_{nn}^k).$$

Motivating example (cont.)

- Suppose that A is not diagonal but it is similar to some diagonal matrix D, i.e.,

$$A = PDP^{-1}.$$

- Then we have that

$$A^k = \underbrace{(PD\overbrace{P^{-1})(P}^{=I}D\overbrace{P^{-1})(P}^{=I}DP^{-1})\dots(PDP^{-1})}_{k \text{ times}} = PD^kP^{-1}.$$

- Since D is diagonal, calculating D^k is straightforward and there are only left two matrix multiplications: one on the left by P and one on the right by P^{-1}.
- It costs to obtain P, D, and P^{-1} but once we have these three matrices, we can do a lot of operations in a much cheaper way.

Questions

- Are all matrices similar to diagonal matrices?
- If a matrix is similar to a diagonal matrix, how do we find P and D?

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors

Definition and computation

Eigenvalues and eigenvectors

Sometimes matrix multiplication behaves like scalar multiplication.

Definition 3.3.1 of Johnston (2021)

Suppose A is a square matrix. A non-zero vector v is called an *eigenvector* of A if there is a scalar λ such that

$$Av = \lambda v.$$

Such a scalar λ is called the *eigenvalue* of A corresponding to v.

- If v is an eigenvector, so is cv with $c \neq 0$. Pay attention that 0 is not an eigenvector according to the definition.
- Recommended to watch *Eigenvectors and eigenvalues* of 3Blue1Brown.

Diagonal matrices

If $A = \mathrm{diag}(a_{11}, \ldots, a_{nn})$, then $e_1, \ldots, e_n$ are the eigenvectors of A and $a_{11}, \ldots, a_{nn}$ are the corresponding eigenvalues since

$$Ae_i = a_{ii}e_i$$

for $i = 1, \ldots, n$.

A two-step procedure: step 1

- λ is an eigenvalue of A if and only if there exists a non-zero vector v such that
$$Av = \lambda v.$$
- We have that
$$Av = \lambda v \iff Av - \lambda v = 0 \iff (A - \lambda I)v = 0.$$
- When does a homogeneous systems of linear equations have a non-zero solution?
- A scalar λ is an eigenvalue of A if and only if $A - \lambda I$ is not invertible.
- $A - \lambda I$ is not invertible if and only if $\det(A - \lambda I) = 0$.

Blackboard example: calculating eigenvalues

Compute the eigenvalues of

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}.$$

A two-step procedure: step 2

- Once we know the eigenvalues of a matrix, the eigenvectors can be found by solving the linear system $Av = \lambda v$ for v.
- This is equivalent to solving the system of linear equations

$$(A - \lambda I)v = 0$$

  (i.e., computing the null space of the matrix $A - \lambda I$).
- A non-zero vector v is an eigenvector of A with corresponding eigenvalue λ if and only if $v \in \mathrm{null}(A - \lambda I)$.

Blackboard example (cont.)

Compute the eigenvectors of

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$$

using the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 6$ from the earlier computations.

Eigenvalues and eigenvectors

Characteristic polynomial and algebraic multiplicity

The characteristic polynomial

$\det(A - \lambda I)$ is a polynomial in λ which is set to 0 and solved for λ when eigenvalues are computed.

Definition 3.3.2 of Johnston (2021)

Suppose A is a square matrix. The function $p_A : \mathbb{R} \to \mathbb{R}$ defined by

$$p_A(\lambda) = \det(A - \lambda I)$$

is called the *characteristic polynomial* of A.

For example, the characteristic polynomial of

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$$

is $p_A(\lambda) = \lambda^2 - 5\lambda - 6$.

Properties of the characteristical polynomial

- The characteristic polynomial of an $n \times n$ matrix has degree n.
- Every degree-n polynomial has at most n distinct roots.
- Every $n \times n$ matrix has at most n distinct eigenvalues.

Definition 3.3.3 of Johnston (2021)

Suppose that $A \in \mathbb{R}^{n \times n}$ with eigenvalue λ. The *algebraic multiplicity* of λ is its multiplicity as a root of A's characteristic polynomial.

- Algebraic multiplicity tells us "how many times" the particular eigenvalue occurs as a root of the characteristic polynomial.
- Roots with multiplicity 2 or greater are sometimes called "repeated roots".

Blackboard example: eigenvalues and algebraic multiplicities

Compute the eigenvalues, and their algebraic multiplicities, of the matrix

$$A = \begin{bmatrix} 2 & 0 & -3 \\ 1 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix}.$$

References

Johnston, Nathaniel (2021). ***Introduction to linear and matrix algebra.*** Springer Cham.