Lecture2: Linear Algebra

Guoxuan Ma[1]

UIBE Math Camp, 2022

Some Notations

Let A be a matrix of M rows and N columns.

$$A_{M \times N} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MN} \end{bmatrix}$$

The element in row i and column j is $[a_{ij}]$
It can also be written as a collection of N M-sized column vectors

$$\left[a^1, a^2, ..., a^N\right], \text{ where } a^i \in \mathbb{R}^M$$

Basic Definitions

- Real numbers can be seen as a 1×1 matrix.
- Matrices with as many rows as columns $m = n$ are called square matrices.
- The zero matrix of M_{MN} is the matrix with all entries equal to zero.
- A square matrix A is diagonal if all its non-diagonal elements are zero: $a_{ij} = 0$ for all i, j such that $i \neq j$. We note $A = diag(a_{11}, \ldots, a_{nn})$.
- The unit matrix of size n is the square matrix of size n having all its components equal to zero except the diagonal components, equal to 1. It is noted I_n.
- A square matrix A is upper-triangular if all its elements below its diagonal are nil: $a_{ij} = 0$ for all $i > j$.
- A square matrix A is lower-triangular if all its elements above its diagonal are nil: $a_{ij} = 0$ for all $i < j$.

Some Basic Matrix Algebra Operations

- Scalar multiplication

$$\alpha A = [\alpha A_{ij}]$$

- Addition

$$A + B = [a_{ij} + b_{ij}]$$

- Multiplication

$$A_{M\times N} \cdot B_{N\times Q} = D_{M\times Q} = [d_{ij}]$$

$$\text{where } d_{ij} = \sum_{k=1}^{N} a_{ik} b_{kj}$$

Example

Multiplying Matrices

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \\ c_7 & c_8 & c_9 \end{bmatrix}$$

Laws of Matrix Algebra

1. $A+B=B+A$
2. $(A+B)+C=A+(B+C)$
3. $(AB)C=A(BC)$
4. $A(B+C)=AB+AC$
5. $(A+B)C=AC+BC$

Transpose

- Definition

$$A_{M\times N} = [a_{ij}]$$
$$A^T_{N\times M} = \left[a^T_{ji}\right] = [a_{ij}] = A_{M\times N}$$

Theorem

1. $(A+B)^T = A^T + B^T$
2. $(AB)^T = B^T A^T$
3. $(\alpha A)^T = \alpha A^T$
4. $(A^{-1})^T = (A^T)^{-1}$ (provided the inverse exists)

- Obviously, if we apply the transpose operator twice, we end up back on A: $(A^T)^T = A$. Note that a row vector is the transpose of a column vector.

Trace

Definition Given $A_{N\times N'}$

$$tr(A) = \sum_{i=1}^{N} a_{ii}$$

Theorem

Given: $A_{N\times N}, B_{N\times N}$

1. $tr(A+B) = tr(A) + tr(B)$
2. $tr(AB) = tr(BA)$
3. $tr(A^T) = tr(A)$

Determinants

For a square matrix A, its determinant, denoted as $\det(A)$, is an element defined inductively in the following way:
(1) For a 1×1 matrix $A = a_{11}$, defined its determinatn are $\det(A) := a_{11}$
(2) For an $n \times n$ matrix where $n \geq 2$,defin its determinant as

$$\det(A) := \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{-1,-j})$$

where $A_{-i,-j}$ is the matrix A with the i-th row and j-th column eliminated.

Example

For a 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

The determinant is

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

For a 3×3 matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

The determinant is

$$\begin{aligned} \det(A) = a_{11} \det\left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \right) - a_{12} \det\left(\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \right) \\ + a_{13} \det\left(\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right) \end{aligned}$$

Determinants

Given $A_{M\times N}$, $|A|$ denotes the determinant of A.
Definition (i,j)-th minor of A

$$A_{ij} : (M-1)\times(N-1)$$

Definition (i,j)- th cofactor of A

$$c_{ij}(A) = (-1)^{(i+j)}|A_{ij}|$$

and

$$|A| = \sum_{k=1}^{N} a_{ij}c_{ij}(A)$$

expansion along i-th row $(k=j)$ or j-th col $(k=i)$

Example

Calculate the determinants of matrix A

$$A = \begin{bmatrix} 2 & 6 \\ 5 & 4 \end{bmatrix}$$

we can get

$$|A| = 8 - 30 = -22$$

Property of Determinants

- $|A| = |A^T|$
- $|AB| = |A||B|$
- $|I_N| = 1$
- $|\lambda A| = \lambda^N |A|$
- $|A| = 0$ if A has a row or col of o's
- If we multiply a row(col) of A by λ to get $\hat{A}$, $|\hat{A}| = \lambda |A|$
- If we add a multiple of a row (col) to another row (col) to get $\hat{A}$, $|\hat{A}| = |A|$

Inverse

Definition Given $A_{N\times N}$, its inverse, denoted A^{-1}, is the $N \times N$ matrix such that

$$A^{-1}A = AA^{-1} = I_N$$

Theorem

1. A^{-1} exists $\Leftrightarrow |A| \neq 0$.
2. If A^{-1} exists, it is unique.
3. We say that A is non-singular if A^{-1} exists.

Properties of the Inverse

- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $|A^{-1}| = \frac{1}{|A|}$
- $(A^T)^{-1} = (A^{-1})^T$

Computation of the Inverse

Definition the Adjoint of an $N \times N$ matrix is given by

$$Adj(A) = [C(A)]^T$$

where

$$C(A) = [C_{ij}(A)]$$

$C(A)$ is a matrix whose element i, j is the (i, j)-th cofactor of A

$$c_{ij}(A) = (-1)^{(i+j)} |A_{ij}|$$

The inverse of A, A^{-1} is given by:

$$A^{-1} = \frac{1}{|A|} Adj(A)$$

Inverse Example

Compute the Inverse of A

$$A = \begin{bmatrix} 2 & 6 \\ 5 & 4 \end{bmatrix}$$

We get

$$A^{-1} = \begin{bmatrix} -2/11 & 3/11 \\ 5/22 & -1/11 \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The inverse of a matrix is found using the following formula:

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$$

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ \\ \begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} \\ \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{11} \\ a_{32} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$

Adjoint Matrix of A

$$adj\,\mathrm{A} = \begin{bmatrix} +\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & +\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & +\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ +\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & +\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$

Linear Dependence and Independence

Definition Non-zero vector $\{x^1, \dots, x^k\} \in \mathbb{R}^N$ are linearly dependent if $\exists \alpha_1, \dots, \alpha_k$ not all zero such that

$$O_N = \sum_{i=1}^{k} \alpha_i x^i$$

where O_N is the N-dimensional zero vector.
Definition $\{x^1, \dots, x^k\} \in \mathbb{R}^N$ are linearly independent if

$$O_N = \sum_{i=1}^{k} \alpha_i x^i \Rightarrow \alpha_i = 0 \ \forall i = 1, \dots, K$$

Rank

Definition The column (row) rank of $A_{M\times N}$ is the max number of linearly

Theorem

independent cols (rows).

$$\text{col rank } (A_{M\times N}) \leq N$$
$$\text{row rank } (A_{M\times N}) \leq M$$

$$\text{col rank}(A) = \text{row rank}(A) = \text{rank}(A)$$

Lemma

- $rank(A) \leq \min\{N, M\}$
- $rank(A^T) = rank(A)$

Definition

$A_{M\times N},\ M \leq N$ is of full rank iff $rank(A) = M$

A square matrix A of size n is invertible iff $rank(A) = n$.

Example

In a 2D setting, if

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2} \text{ let } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^{2\times 1},$$

Then $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y = Ax = \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{pmatrix}$

We see that because the rank of matrix A is 1 and $y_2 = 2y_1$ which is nothing but a line passing through the origin in the plane

Systems of Linear Equations

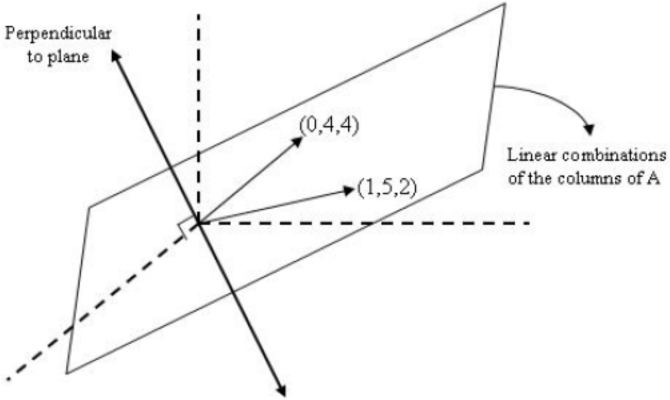
Consider the following system of linear equations in x:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

where a_{ij} 's and b_i's are all elements of $\mathbb{R}$, and the unknowns $x_1, \ldots, x_n$ also take values in $\mathbb{R}$.
In matrix format, we have

$$A_{M\times N} \cdot x_{N\times 1} = b_{M\times 1}$$

Geometric Intuition



$Ax = b$ can be solved iff b lies in the plane that is spanned by the column vectors of A.

Connection to Rank

- You cannot solve a linear system when it is rank deficient. The matrix A maps x to y is neither onto nor one to one.
- The rank of a matrix denotes the information content of the matrix.
- In general, if we know that a matrix $A \in \mathbb{R}^{m \times n}$ is of rank p, then we can write A as $U \cdot V^T$ where $U \in \mathbb{R}^{m \times p}$ and is of rank p and $V \in \mathbb{R}^{n \times p}$ and is of rank p.
- From solving a linear system point of view, when the square matrix is rank deficient, it means that we do not have complete information about the system, ergo we cannot solve the system.

Example

Solving the system of linear equations

$$\begin{aligned} -12x_1 + 9x_2 &= 7 \\ 3x_1 - 4x_2 &= 2 \end{aligned}$$

Eigenvalues and Eigenvectors

Definition

Let A be an $m \times n$ matrix over $\mathbb{C}$. A scalar $\lambda \in \mathbb{C}$ is said to be an eigenvalue of A iff $\exists x \in \mathbb{C}^n \setminus \{0\}$ s.t. $Ax = \lambda x$. A vector $x \in \mathbb{C}^n \setminus \{0\}$ is said to be an eigenvector of A iff $\lambda \in \mathbb{C}$ s.t. $Ax = \lambda x$.

Proposition

$\lambda \in \mathbb{C}$ is an eigenvalue of A iff $\det(\lambda I_n - A) = 0$.

By definition, $Ax = \lambda x$ has nonzero solution is equivalent to $(A - \lambda I_n)x = 0$ has a nonzero solution.

Theorem

Let $P : \mathbb{C} \to \mathbb{C}$ be a polynomial of degree n, i.e.
$P(\lambda) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$, *where* $c_k \in \mathbb{C}$ *for any* $k = 0, 1, \ldots, n$ *and* $c_n \neq 0$. *Then P has exactly n roots in $\mathbb{C}$, counted with multiplicity. That is, there exists* $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ *s.t.*

$$P(\lambda) = c_n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

Therefore, we can obtain all eigenvalues of A by setting the characteristic polynomial of A to 0 and solving for all its roots.

Example

Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix}$
First, we can find the eigenvalues of A by solving

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

This gives $\lambda^2 + \lambda - 6 = 0$, which we find $\lambda_1 = 2$ and $\lambda_2 = -3$.
Plug in to the

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

for each eigenvalue and by normalization, we can find

$$X_1 = \begin{bmatrix} 2 \\ 7 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Diagonalization

Let A be an $n \times n$ matrix over $\mathbb{C}$ and $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{C}$ are eigenvalues of A. Then $\det(A) = \lambda_1 \lambda_2 ... \lambda_n$

Definition

Let A be an $n \times n$ matrix over $\mathbb{C}$. The matrix A is diagonalizable in C iff there exists an $n \times n$ invertible matrix P over $\mathbb{C}$ and an $n \times n$ diagonal matrix over $\mathbb{C}$ s.t. $P^{-1}AP = \Lambda$.

Let A be an $n \times n$ matrix over $\mathbb{C}$. Then A is diagonalizable in $\mathbb{C}$ iff A has n linearly independent eigenvectors.

Example

Following the previous question, given $A = \begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix}$, let do the Diagonalization.

From the previous question, we already learn that the eigenvalues are $\lambda_1 = 2$, and $\lambda_2 = -3$, with the corresponding eigenvectors

$$C = \begin{bmatrix} 2 & 1 \\ 7 & 1 \end{bmatrix}$$

So it is straightforward to have

$$A = C \times \Lambda \times C^{-1}$$

where $\Lambda = Diag(2, -3)$.