Mathematical Methods

Giorgio Grigolo

November 17, 2021

Lecture 1: Matrix definitions and operations

Definition 1. An $m \times n$ matrix is a rectangular array of numbers, arranged in m rows and n columns. The elements of said matrix are called **entries** and the expression $m \times n$ denotes its **size**.

$$\mathbf{A} = (a_{ij})_{m\times n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

Definition 2. An $n \times n$ matrix is called a **square matrix** of order n.

The entries $\{\, a_{11}, a_{22}, \ldots, a_{nn} \,\}$ form the **main diagonal**.

Definition 3. Two matrices are equal if they have the **same dimensions** and all their corresponding entries are equal.

$$a_{ij} = b_{ij},\ \forall\ i,\ j \in \mathbb{N}$$

Let $\mathbf{A} = \big(a_{ij}\big)_{m\times n}$ and $\mathbf{B} = \big(b_{ij}\big)_{m\times n}$, then

$$\begin{aligned} \mathbf{A}+\mathbf{B} &= \big(a_{ij}+b_{ij}\big)_{m\times n} \\ \mathbf{A}-\mathbf{B} &= \big(a_{ij}-b_{ij}\big)_{m\times n} \\ \lambda\mathbf{A} &= \big(\lambda a_{ij}\big)_{m\times n} \\ \mathbf{AB} &= \big(c_{ij}\big)_{m\times n} \end{aligned}$$

where $c_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj}$.

Note. In general, for two matrices **A** and **B**, $\mathbf{AB} \neq \mathbf{BA}$, because

- **AB** could be defined but **BA** is not.
- both are defined but **AB** and **BA** have different size.
- both might be defined and have the same size, but $\mathbf{AB} \neq \mathbf{BA}$.

Lecture 2: Gaussian-Jordan Elimination

A set of linear equations in the variables $x_1, x_2, \cdots, x_n$ is called a system of linear equations (linear system). A system of m linear equations in n unknowns can be written as:

$$\left.\begin{array}{ccccccc} a_{11}x_1 &+& a_{12}x_2 &+\cdots+& a_{1n}x_n &=& b_1 \\ a_{21}x_1 &+& a_{22}x_2 &+\cdots+& a_{2n}x_n &=& b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 &+& a_{m2}x_2 &+\cdots+& a_{mn}x_n &=& b_m \end{array}\right\}$$

A linear system can be solved using the augmented matrix

$$\begin{pmatrix} A & B \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}$$

A solution of a linear sequence is a sequence of numbers $s_1, s_2, \cdots, s_n$ s.t. $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$ is a solution of every equation.

Solutions can be of 3 types:

- Infinite
- Unique
- Non-existent

An example of **non-existent** (*inconsistent*) system of equation would be

$$\left.\begin{array}{ll} x+y & =2 \\ 2x+2y & =6 \end{array}\right\}$$

To solve consistent systems of linear equations one must use the **elementary row operations** to reduce the augmented matrix $\begin{pmatrix} A & B \end{pmatrix}$ into **row echelon form**.

The following algorithm (*also referred to as Gaussian Elimination*) must be followed to reduce an $m \times n$ matrix into a row echelon form one.

1. If a row has all entries 0, then it must be placed at the bottom.
2. If a row does contain an entry which is not 0, then the first non-zero entry must be 1 (*also referred to as the leading one*).
3. In any two successive rows, the bottom one must have the leading 1 furhter to the right than that of the higher.

If a reduced echelon form matrix is desired, it must be in row echelon form and have every entry of each column which contains a leading one (*except the leading one*), be 0. The process by which a reduced echelon form matrix is obtained is called *Gaussian-Jordan Elimination.*

Lecture 3: Rank of a Matrix

Definition 4. The rank of a matrix, denoted by $\operatorname{rank}(\mathbf{A})$, is equal to the number of non-zero rows in a row echelon form of $\mathbf{A}$

Theorem 1. Let $\mathbf{AX} = \mathbf{B}$ be a linear system of m linear equations in n unknowns with augmented matrix $\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$, then

- the system has a solution if and only if $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$.
- the system has a unique solution if and only if $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} = n$.

Note. The rank of a matrix augmented with another cannot be smaller than the original non-augmented matrix and thus

$$\operatorname{rank}(\mathbf{A}) \neq \operatorname{rank}\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} \Leftrightarrow \operatorname{rank}(\mathbf{A}) < \operatorname{rank}\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$$

For which values of a does the following system have a unique solution? For which pairs of a, b does the system have more than one solution?

$$\left.\begin{aligned} x - 2y &= 1 \\ x - y + az &= 2 \\ ay + 9z &= b \end{aligned}\right\}$$

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 & 1 \\ 1 & -1 & a & 2 \\ 0 & a & 9 & b \end{pmatrix} \underset{\sim}{R} \begin{pmatrix} 1 & -2 & 0 & 1 \\ 0 & 1 & 9 & 1 \\ 0 & 0 & 9-a^2 & b-a \end{pmatrix}$$

The solution is unique if and only if $a \neq \pm 3$. Thus, $\operatorname{rank}(\mathbf{A}) = 3$. The solutions are infinite if and only if $a = \pm 3$ and $b - a = 0$.

$$(a, b) = (\pm 3, \pm 3) \Rightarrow \operatorname{rank}(\mathbf{A}) = 3$$

If a system of n equations in n unknowns (*also referred to as a square system*) has a unique solution, then the solution can be found by using the inverse of the coefficient matrix.

$$\left.\begin{aligned} \mathbf{AX} &= \mathbf{B} \\ A \text{ is } n &\times n \\ \text{The system has 1 solution} \end{aligned}\right\} \mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

Note. An $n \times n$ matrix $\mathbf{A}$ is invertible if and only if $\operatorname{rank}(\mathbf{A}) = n$.

Lecture 4: Homogenous Linear Equations

A system of m equations in n variables $x_1, x_2, \cdots x_n$ is homogeneous if it is in the form

$$\mathbf{AX} = \mathbf{0}.$$

Such a system always has one solution (*the trivial solution*), namely $x_1 = 0, x_2 = 0, \cdots, x_n = 0$. Such a system might also have an infinite set of solutions in the possibility that $\operatorname{rank}(\mathbf{A}) < n$.

Remark. If the number of equations (m) is less than the number of variables (n) *i.e.* $m < n$, then the system has a non trivial solution.

Lecture 5: Determinants

Definitions

Let $\mathbf{A}$ be a square matrix. The scalar $\det(A)$ or $|\mathbf{A}|$ is called the **determinant** of $\mathbf{A}$.

For a 2×2 square matrix, the determinant is given by $a_{11}a_{22} - a_{12}a_{21}$. The plot thickens very quickly as we try to determine the determinant of larger $n \times n$ matrices.

The process of evaluating the determinant of an $n \times n$ matrix is described by the following steps:

1. The **minor** $\mathbf{M_{jk}}$ is the matrix obtained from a square matrix when omitting the j^{th} row and the k^{th} column.
2. The scalar quantity referred to as the **cofactor** $\alpha_{jk} = (-1)^{j+k}\det(\mathbf{M}_{jk})$

The determinant of an $n \times n$ matrix $\mathbf{A}$, when choosing the first row as the starting point is therefore given by

$$\begin{aligned}\det(\mathbf{A}) &= a_{11}\mathbf{A}_{11} + a_{12}\mathbf{A}_{12} + \cdots + a_{1n}\mathbf{A}_{1n} \\ &= \sum_{i=1}^{n} a_{1r}\mathbf{A}_{1r}\end{aligned}$$

The same value of the determinant is obtained if the first point of reference is chosen to be any other row or column. To minimize computation complexity, the row or column with the most entries equal to 0 should be chosen with the result of reducing the maximum number of terms in the sequence to 0.

Properties of Determinants

P1. $\det(\mathbf{A}) = \det(\mathbf{A}^T)$

P2. If $\mathbf{A}$ is an $n \times n$ triangular matrix, then $\det(\mathbf{A}) = a_{11}a_{22}\cdots a_{nn}$.

P3. The interchanging of any two consecutive rows only alters the sign of of the determinant without affecting its magnitude.

P4. If any two rows consecutive rows or columns of an $n \times n$ matrix $\mathbf{A}$ are equal, then $\det(\mathbf{A}) = 0$.

Theorem 2. If all elements of one row are multiplied by a constant k, then the value of the determinant is also multiplied by k.

Proof. Let $\mathbf{K}$ be the general $n \times n$ matrix $\mathbf{A}$ but with all entries in the first row multiplied by k. Then,

$$\begin{aligned}
|\mathbf{K}| &= \begin{vmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\
&= ka_{11}A_{11} + ka_{12}A_{12} + \cdots + ka_{1n}A_{1n} \\
&= k(a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}) \\
&= k\,|\mathbf{A}|
\end{aligned}$$

□

Corollary 1. If any two rows of an $n \times n$ matrix $\mathbf{A}$ are multiples of the other (*are linearlly dependent*), then the $|\mathbf{A}| = 0$.

Corollary 2. If $\mathbf{A}$ is an $n \times n$ matrix, then $\det(k\mathbf{A}) = k^n \det(\mathbf{A})$.

Theorem 3. If a distinct scalar α_{ij} is added to all the entries of a row of an $n \times n$ matrix $\mathbf{A}$, then $\det(\mathbf{A})$ is equal to

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \dots \\ a_{n_1} & a_{n_2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_1} & a_{n_2} & \cdots & a_{nn} \end{vmatrix}.$$

Proof.

$$\text{Consider} \quad \begin{vmatrix} a_{11}+\alpha_{11} & a_{12}+\alpha_{12} & \cdots & a_{1n}+\alpha_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \dots \\ a_{n_1} & a_{n_2} & \cdots & a_{nn} \end{vmatrix}$$

$$\begin{aligned}
&= (a_{11}+\alpha_{11})\,\mathbf{A}_{11} + (a_{12}+\alpha_{12})\,A_{12} + \cdots + (a_{1n}+\alpha_{1n})\,\mathbf{A}_{1n} \\
&= (a_{11}\mathbf{A}_{11} + a_{12}\mathbf{A}_{12} + \cdots + a_{1n}\mathbf{A}_{1n}) + (\alpha_{11}\mathbf{A}_{11} + \alpha_{12}\mathbf{A}_{12} + \cdots + \alpha_{1n}\mathbf{A}_{1n})
\end{aligned}$$

□

Theorem 4. The value of a determinant is unchanged if we add to the entries of any row the same multiple of the corresponding entries of another row.

$$\text{If } |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \text{ then } \begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} & \cdots & a_{1n}+ka_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = |\mathbf{A}|$$

Proof. Let $\mathbf{K} = \begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} & \cdots & a_{1n}+ka_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$. Then, by Theorem **??**,

$$|\mathbf{K}| = |\mathbf{A}| + \begin{vmatrix} ka_{21} & ka_{22} & \cdots & ka_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Then, by Theorem **??** and P4, $\begin{vmatrix} ka_{21} & ka_{22} & \cdots & ka_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$

Therefore, $|\mathbf{K}| = |\mathbf{A}|$. □

Theorem 5. The determinant of the product of two matrices is equal to the product of the two determinants.

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$$

Proof. This proof is left as an exercise to the reader for the 2×2 matrix case. □

Theorem 6. If $\mathbf{A}^{-1}$ is invertible, then $\left|\mathbf{A^{-1}}\right| = \frac{1}{\det(\mathbf{A})}$

Proof. $\mathbf{A}$ is invertible. Thus

$$\begin{aligned} \exists\ \mathbf{B} \text{ s.t. } \mathbf{AB} &= I \\ \det(\mathbf{AB}) &= \det(I) \\ \det(\mathbf{A}) \cdot \det(\mathbf{B}) &= 1 \\ \det(\mathbf{B}) &= \frac{1}{\det(\mathbf{A})} \end{aligned}$$

$\square$

Remark. An $n \times n$ matrix is invertible $\Leftrightarrow \det(\mathbf{A}) \neq 0$

Lecture 6: Eigenvalues

Definition 5. Let $\mathbf{A}$ be an $n \times n$ matrix. A non-zero n-vector $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

is called an **eigenvector** of $\mathbf{A}$ with corresponding **eigenvalue** λ.

To find the eigenvalues, consider $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. After applying some algebraic manipulations, particularly the factorisation of $\mathbf{x}$, which involves the multiplication of λ and the identity matrix, we get:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$$

Above equation is a homogenous system of equations which must ahve a solution other than $\mathbf{x} = 0$. The eigenvalues of any $n \times n$ matrix $\mathbf{A}$ are the values of λ that satisfy the following equation, which is referred to as the *characteristic equation*.

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

Note. For an $n \times n$ matrix, the above equation turns out to be just a polynomial of degree n in terms of λ. As expected, its roots might be real or complex as well as distinct or equal.

Lecture 7: Eigenvectors

To find the *eigenvectors* of an $n \times n$ matrix, or rather, a parametrization of the family of eigenvectors for the given matrix, we substitute the previously obtained eigenvalues into the characteristic equation:

$$\left.\begin{aligned} (\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} &= \mathbf{0} \\ (\mathbf{A} - \lambda_2\mathbf{I})\mathbf{x} &= \mathbf{0} \\ &\vdots \\ (\mathbf{A} - \lambda_n\mathbf{I})\mathbf{x} &= \mathbf{0} \end{aligned}\right\}$$

Ex. Consider the matrix $\mathbf{A} = \left(\begin{smallmatrix} 3 & 0 \\ 8 & -1 \end{smallmatrix}\right)$ whose eigenvalues are -1 and 3.

$$\begin{pmatrix} 3 - \lambda_1 & 0 \\ 8 & -1 - \lambda_1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = 0$$

Letting $\lambda_1 = -1$ and substituting into the characterstic equation we get:

$$\left(\begin{array}{cc|c} 4 & 0 & 0 \\ 8 & 0 & 0 \end{array}\right) \underset{\sim}{R} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

$$\Rightarrow \quad \left.\begin{aligned} x_{11} &= 0 \\ x_{12} &= t \neq 0 \end{aligned}\right\} \mathbf{x}_1 = t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Letting $\lambda_2 = 3$ and substituting into the characterstic equation we get:

$$\left(\begin{array}{cc|c} 0 & 0 & 0 \\ 8 & -4 & 0 \end{array}\right) \underset{\sim}{R} \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

$$\Rightarrow \left.\begin{array}{l} x_{11} = r \neq 0 \\ 2x_{11} = x_{12} = 2r \end{array}\right\} \mathbf{x}_2 = t \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Thus, $\begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ has eigenvectors $\mathbf{x}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Theorem 7. If $\mathbf{A}$ is an $n \times n$ matrix, then the following are equivalent:

- λ is an eigenvalue of $\mathbf{A}$.
- there is a non-zero vector $\mathbf{x}$ of dimension n such that $\mathbf{Ax} = \lambda\mathbf{x}$.
- λ is the solution of $\det(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$.
- the system of equations $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$ has an infinite number of non-trivial solutions.

Note. The complete solution set of the homogenous system of equations $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$ is called the **eigenspace** of $\mathbf{A}$ corresponding to the eigenvalue λ and is denoted by E_λ. The elements of the set of particular solutions obtained by solving the the above system of equations using Gaussian elimination and setting each independent variable to 1 and all the other variables to 0 are called the **fundamental eigenvectors**.

Lecture 8: Diagonalization

An $n \times n$ matrix $\mathbf{A}$ is diagonalisable if it is "similar" to a diagonal matrix $\mathbf{D}$, that is, there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.

In this case, $\mathbf{D}$ is the diagonal matrices whose entries on the main diagonal are the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ of the initial matrix $\mathbf{A}$. $\mathbf{P}$ is the square matrix whose column vectors are the eigenvectors of the initial matrix $\mathbf{A}$.

Definition 6. An $n \times n$ matrix $\mathbf{A}$ is diagonalisable if the number of fundamental eigenvectors is equal to its dimension.

Recall $\mathbf{A} = \left(\begin{smallmatrix} 3 & 2 \\ -1 & 0 \end{smallmatrix}\right)$, whose eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$; and whose fundamental eigenvectors are $\left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right)$.

Ex. Thus,

$$\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad \begin{array}{l} \mathbf{P} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \\ \mathbf{P}^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \end{array}$$

Problem. Find $\mathbf{A}^5$, where $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$. But $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \Rightarrow \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$. Then,

$$\begin{aligned}
\mathbf{A}^5 &= (\mathbf{P}\mathbf{D}\mathbf{P}^{-1})^5 \\
&= \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \\
&= \mathbf{P}\mathbf{D}^5\mathbf{P}^{-1}
\end{aligned}$$

Note. If $\mathbf{A}$ is diagonalisable, then $\mathbf{A}^n = \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}$