

Matrices

Paolo Bettelini

Contents

1	Definition	2
2	Linear Transformation	2
2.1	Definition	2
2.2	The basis	2
3	Operations with matrices	3
3.1	Matrix multiplication	3
3.2	Matrix transpose	3
3.3	Matrix inverse	4
4	Types of matrices	5
4.1	Identity matrix	5
4.2	Matrix adjoint	5
4.3	Hermitian Matrix	5
4.4	Vandermonde matrix	5
4.5	Determinants	6
5	Column space	7
6	Rank of a matrix	7
7	Null space	7
8	System of linear equations	8
8.1	Interpretation	8
8.2	Cramer's rule	8
9	Change of basis	9
9.1	Change of basis of a vector	9
9.2	Change of basis of a linear transformation	9
10	Eigenvalues and eigenvectors	9
11	Rotational matrix	10

1 Definition

A matrix is a rectangular array of elements. Each matrix has a particular size expressed as $n \times m$. A generic matrix looks like the following

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}$$

2 Linear Transformation

2.1 Definition

Multiplying a vector by a matrix produces another vector. This is a linear transformation. A linear transformation T moves a vector around the vector space and/or to another space.

Applying a *linear* transformation doesn't move the origin of the space, and every line is kept linear. This operation can stretch and rotate every vector on the plane. The matrix contains the information about this transformation.

$$\mathcal{V} \xrightarrow{T} \mathcal{W}$$

The linear transformation has the following proprieties:

1. $T(c\vec{a}) = cT\vec{a}$
2. $T(\vec{a} + \vec{b}) = T\vec{a} + T\vec{b}$

where c is a scalar of the vector space.

2.2 The basis

Given a basis for a vector space

$$\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$$

we can expand a vector \vec{a} along this basis

$$\vec{a} = \sum_{k=1}^n \alpha_k \mathcal{B}_k$$

We then apply a transformation T to the vector \vec{a} and use the proprieties of linear transformations

$$T\vec{a} = T \sum_{k=1}^n \alpha_k \mathcal{B}_k = \sum_{k=1}^n \alpha_k T\mathcal{B}_k$$

This tells us that we only need $T\mathcal{B} = \{T\vec{b}_1, T\vec{b}_2, \dots, T\vec{b}_n\}$ to determine T applied to any vector on the vector space. Any transformation T is completely described by \mathcal{B} and $T\mathcal{B}$.

Each column of a matrix is indeed the result of applying its transformation to the corresponding vector in the basis. Intuitively, this is given by the fact that we only need to know where the vectors of the basis end up after the transformation in order to represent the whole information.

3 Operations with matrices

3.1 Matrix multiplication

Matrix multiplication forms a composition of multiple linear transformations. This operation is not a commutative operation.

$$AB \neq BA$$

However it is an associative operation.

$$A(BC) = (AB)C$$

Matrix multiplication can be computed only if the number of columns of the first matrix is equal to the number of rows of the other matrix.

This means that we can compute the product of a $n \times m$ matrix with a $i \times j$ matrix, only if $m = i$. The resulting product is a matrix with n rows and j columns.

$$(n \times m) \cdot (m \times j) = (n \times j)$$

If we multiply a $1 \times n$ matrix by a $n \times 1$ one, the result is a 1×1 matrix or a scalar.

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{k=1}^n a_k b_k = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

This operation is pretty much the same as the dot product in vectors.

For bigger matrices, we apply the same ‘dot product’ between every row and every column. Here is an example:

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix} \cdot \begin{bmatrix} d_1 & \mathbf{e}_1 \\ d_2 & \mathbf{e}_2 \\ d_3 & \mathbf{e}_3 \\ d_4 & \mathbf{e}_4 \end{bmatrix} = \begin{bmatrix} a \cdot d & a \cdot \mathbf{e} \\ b \cdot d & \mathbf{b} \cdot \mathbf{e} \\ c \cdot d & c \cdot \mathbf{e} \end{bmatrix}$$

Where $a \cdot b$ denotes the ‘dot product’ between a row and a column.

3.2 Matrix transpose

The transpose operator is analogous to that of vectors

$$(a_1 \ a_2 \ \cdots \ a_n)^t = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

This is actually a special case, where we apply the transpose operator to a matrix column.

The transpose of a generic $n \times m$ matrix results in a $m \times n$ one.

$$(A_{ij})^t = A_{ji}$$

The transpose of a matrix is just a flipped version of the original matrix. We can transpose a matrix by switching its rows with its columns. The original rows become the new columns and the original columns become the new rows.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}^t = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix}$$

3.3 Matrix inverse

An $n \times n$ matrix A is invertible (also nonsingular or nondegenerate), if there exists an $n \times n$ matrix A^{-1} such that

$$AA^{-1} = I_n$$

If $\det(A) = 0$, meaning that the matrix collapses space into a lower dimension, thus resulting in a loss of information, the matrix is not invertible. Otherwise, the matrix has an inverse and it is unique.

This is equivalent to saying that if $A\vec{v} = 0$ for some non-zero vector \vec{v} , then A has no inverse.

4 Types of matrices

4.1 Identity matrix

The identity matrix is a square matrix with ones on the main diagonal and zeros elsewhere. This matrix is denoted as I_n where n is the length of the side.

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

When I_n is applied to a matrix or vector, the matrix or vector remains the same.

4.2 Matrix adjoint

Given a complex matrix A we form the transpose and then take the complex conjugate of every element in it.

$$A^\dagger \equiv (A^t)^*$$

4.3 Hermitian Matrix

A matrix M is called Hermitian if it is equal to its adjoint.

$$M^\dagger = M$$

4.4 Vandermonde matrix

A Vandermonde matrix is a matrix with the terms of a geometric progression in each row or column. Here is an $m \times n$ matrix.

$$V = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_m & a_m^2 & \cdots & a_m^{n-1} \end{bmatrix}$$

4.5 Determinants

Every square matrix has a scalar called the determinant.

The determinant is the factor by which the area of any shape (or volume and so on) changes after the transformation is applied. If the transformation inverts the orientation of the space, the determinant is negative.

To compute the determinant of a matrix, we must define the minor of a matrix element.

The minor of a matrix element is the part of the matrix remaining after excluding the row and the column containing that particular element.

$$\text{minor of } b = \begin{bmatrix} \cancel{a} & \cancel{b} & \cancel{e} \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{vmatrix} d & f \\ g & i \end{vmatrix}$$

To compute the determinant of a matrix, for each element of a row or column, we multiply the element by the determinant of its minor. We then sum all of these values, but every even term has a negative sign. This is a recursive operation for $n > 2$.

For example, given a 2×2 matrix A

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

its determinant is defined as

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \equiv ad - bc$$

For the 3×3 case

$$B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$\det(B) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

In general, we can define the determinant of a matrix using the Laplace expansion.

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} A_{ij} \det(\text{minor of } A_{ij})$$

where i is any row.

Instead of expanding the series along a row we could expand it along any column.

The result is always the same no matter what row or what column.

The determinant has some properties, for instance

$$\det(A^t) = \det(A)$$

$$\det(AB) = \det(A) \det(B)$$

5 Column space

The column space (or range or image) of a matrix is the set of all possible vectors that can be generated using the transformation.

6 Rank of a matrix

The rank of a matrix M is the dimension of its column space. A rank n means that every vector after the transformation of M is projected in the n -dimension.

A matrix is full rank if its rank is equal to the number of its columns, meaning the dimension does not change. A matrix is full rank iff $\vec{v} = \vec{0}$ is the only vector such that $M\vec{v} = \vec{0}$.

Note that non-square matrices may still be full rank even if the vectors end up in a lower dimension. If there is no loss of information the matrix is still full rank. If M has dimensions $m \times n$, then it is full rank if $\text{Rank}(M) = \min(n, m)$.

7 Null space

The null space or the kernel of a matrix is the set of all the vectors that land on the origin after the transformation. The kernel is the linear subspace of the domain of the transformation which is mapped to $\vec{0}$.

8 System of linear equations

8.1 Interpretation

A system of linear equations

$$a_1x + b_1y + c_1z = d_1$$

$$a_2x + b_2y + c_2z = d_2$$

$$a_3x + b_3y + c_3z = d_3$$

Can be represented by a matrix multiplication $M\vec{x} = \vec{d}$

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

The geometrical interpretation is to find the vector \vec{x} such that when the matrix M is applied to it, the resulting vector is \vec{d} .

8.2 Cramer's rule

Consider a system of n equations and unknowns

$$A\vec{x} = \vec{b}$$

The solution is given by

$$\vec{x}_i = \frac{\det(A_i)}{\det(A)}$$

where A_i is formed by replacing the i -th column of A by \vec{b} .

9 Change of basis

9.1 Change of basis of a vector

Given a basis \mathcal{B}_1 and \mathcal{B}_2 , there exists a matrix M such that $M\vec{v}$ translates the vector from \mathcal{B}_1 to \mathcal{B}_2 .

The matrix M is formed by the columns of \mathcal{B}_1 but written in \mathcal{B}_2 .

The operation $M^{-1}\vec{v}$ translates the vector from \mathcal{B}_2 to \mathcal{B}_1 .

9.2 Change of basis of a linear transformation

If T_1 is a transformation in \mathcal{B}_1 , we can find another transformation T_2 that is similar to T_1 but works for the basis \mathcal{B}_2 .

Instead of computing $T_1\vec{v}$, we first multiply \vec{v} by the change of basis matrix M , then apply T_1 and finally apply the inverse change of basis matrix M^{-1} .

$$T_2 = M^{-1}T_1M$$

10 Eigenvalues and eigenvectors

The eigenvectors \vec{v} and eigenvalues λ of a matrix M are values such that

$$M\vec{v} = \lambda\vec{v}$$

11 Rotational matrix

The standard rotation matrix in two dimensions has the following form:

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Multiplying a vector $\vec{p} \in \mathbb{R}^2$ by $R(\theta)$ will rotate the point around the origin depending on θ .

$$\vec{p} = \begin{pmatrix} x \\ y \end{pmatrix}$$

We can multiply the rotational matrix by the column vector to obtain the new coordinates

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{bmatrix}$$

Note: the matrix must be positioned on the right of the vector, otherwise the dimensions would no longer be compatible.

The rotation matrices for a three-dimensional point are

$$\begin{aligned} R_x(\theta) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \\ R_y(\theta) &= \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \\ R_z(\theta) &= \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$