

1 Transpose

Transpose can be seen as the behavior of a linear system, that is to say, it contains the properties of linearity. The definition of transpose is as below,

For an arbitrary $n \times m$ matrix A^1 , the transpose of A is,

$$A \longrightarrow A^T = B$$

Where,

$$a_{i,j} = b_{j,i} , \text{ for } i = 1, \dots, n; j = 1, \dots, m$$

By this definition, we see that $A^T = B$ is a $m \times n$ matrix, and simply put, **the rows become columns, and the columns become rows.**

2 Matrix-Vector Product

In this section, we will view the product of matrix and vector with two different aspects. First, given a $n \times m$ matrix A and a vector \mathbf{x} with m rows, then their product can be expressed as,

$$A\mathbf{x} = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1m}x_m \\ \cdot \\ \cdot \\ \cdot \\ a_{n1}x_1 + \dots + a_{nm}x_m \end{bmatrix}$$

In this point of view, we can say that we do the **inner product with the rows and vector**. On the other hand,

¹with n rows and m columns.

$$A\mathbf{x} = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \cdot \\ \cdot \\ \cdot \\ a_{1n} \end{bmatrix} + \dots + x_m \begin{bmatrix} a_{1m} \\ \cdot \\ \cdot \\ \cdot \\ a_{nm} \end{bmatrix}$$

This is the **weighted sums of columns**.

Moreover, we observe that the result of matrix-vector product is **another vector**, in this case, it is a vector **with n rows**. In engineer points of view, the matrix A can represent the linear system, and the vector \mathbf{x} can represent the input. After operating², we get the vector output. In quantum mechanics, we view the operators as matrice, and by this means, we say that the physical state, that is, the ket, after operation(timing by matrix) is another physical state (ket).

Now a question arises, given two $m \times n$ martice A and B , if $A\mathbf{w} = B\mathbf{w}$ for all vectors \mathbf{w} in \mathfrak{R}^n , can we say the two matrice A, B is the same? The proof is simple, we make use of the concept of **standrad basic of \mathfrak{R}^n** , that is, $\mathbf{e}_1, \dots, \mathbf{e}_n$. Therefore, we'll have,

$$A\mathbf{e}_i = B\mathbf{e}_i, \text{ for } i = 1, \dots, n$$

then,

$$A\mathbf{e}_i = \begin{bmatrix} a_{1i} \\ \cdot \\ \cdot \\ \cdot \\ a_{ni} \end{bmatrix}$$

$$B\mathbf{e}_i = \begin{bmatrix} b_{1i} \\ \cdot \\ \cdot \\ \cdot \\ b_{ni} \end{bmatrix}$$

Then we should be sure $A = B$.

²in this case, timing the matrix with the vector input

3 Eigenvalues and Eigenvectors

Eigenvalues are useful when we want to find out the **powers of matrix**. First, we shall talk about what is **Eigenvectors**, it is originated from a concept that **almost every vector will change its direction after the matrix A acting on them**. However, there are some vectors that wouldn't change their direction, that is,

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

Where λ is a real number(or a complex number). And we call λ A's **eigenvalue** and \mathbf{x} A's **eigenvector**.

In **Dirac Notation**, a vector \mathbf{x} can be seen as a **ket**³ and the matrix A can be seen as an **operator**. There Dirac notation is as shown below:

$$A|x\rangle = \lambda|x\rangle$$

4 The Way to Find Eigenvalues and Eigenvectors

As we may see in the following sections, it is worth mentioning the properties with eigenvalue:

m by m matrix should have **m** eigenvalues and **m** eigenvectors.

As for the way to find out eigenvalue, we should use the following formula which would be fully explained in the following context:

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

Example: consider a 2 by 2 matrix A,

$$A = \begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix}$$

and then we shall see $\det(A - \lambda I) = 0$

$$\det \begin{pmatrix} .8 - \lambda & .3 \\ .2 & .7 - \lambda \end{pmatrix} = \lambda^2 - 1.5\lambda + 0.5 = (\lambda - 1)(\lambda - \frac{1}{2})$$

³However, there is a slightly difference between them, that is, the ket emphasizes on its direction but not its magnitude, yet for vector we consider both.

So the two eigenvalues are $\lambda = 1$ and $\lambda = \frac{1}{2}$. For these eigenvalues, we can see that the matrix $A - \lambda I$ becomes *singular matrix*⁴. And since the definition of eigenvector is as below:

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

It is clearly⁵ that

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

So we can conclude that **eigenvectors** are in the **nullspace**⁶ of $A - \lambda I$, that is, we'll have the following equation for the **example**:

$$(A - I)\mathbf{x}_1 = \mathbf{0}$$

If we write it out explicitly,

$$\begin{pmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and then we'll have $0.2x_1 - 0.3y_1 = 0$, thus we can choose the eigenvector \mathbf{x}_1 to be $(.6, .4)$. Similarly, we'll have $\mathbf{x}_2 = (.5, -.5)$. Because of the definition of eigenvalue and eigenvector, we'll have the following properties:

When A is squared, the eigenvectors stay the same, but the eigenvalues are squared.

The fact that **all other vectors are combinations of the eigenvectors**, and the proof needs to be specified. However, for now we shall use this properties and that **Each eigenvector is multiplied by its eigenvalue, when we multiply by A**, and we can accomplish all multiplication of all vectors with matrix A. After a short introduction to eigenvalue and eigenvector, one should know that *special matrix will have special eigenvalue and eigenvector whose patterns and properties are worth studying*.

⁴The definition of singular matrix is that the determination is 0, that is, not invertible

⁵one can check this by writing out the system of equations

⁶Nullspace is the space of solutions of $A\mathbf{x} = \mathbf{0}$, and is denoted by $\mathbf{N}(A)$. One can simply check that we are able to add and multiply without leaving the nullspace, so it is a subspace.

5 Diagonalizing a Matrix

From the above sections, we've already learnt the concept and definition of eigenvector and eigenvalue. In this section, we'll go through the most important application of them, that is, **Diagonalization**.

Suppose the n by n matrix A has n linearly independent⁷ eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Put them into the **columns** of an **eigenvector matrix** S .

Then we would have $S^{-1}AS$ is the **eigenvalue matrix** Λ which is a diagonal matrix with **eigenvalue** λ **on its diagonal**.

The proof is simple, consider A times S :

$$AS = A \begin{pmatrix} \mathbf{x}_1 & \dots & \mathbf{x}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{x}_1 & \dots & \lambda_n \mathbf{x}_n \end{pmatrix}$$

where we make use of the definition of eigenvalue and eigenvector. and then the trick is to **split this matrix AS into S times Λ** .

$$\begin{pmatrix} \lambda_1 \mathbf{x}_1 & \dots & \lambda_n \mathbf{x}_n \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 & \dots & \mathbf{x}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & \dots \\ 0 & \lambda_2 & \dots & 0 & \dots \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix} = S\Lambda$$

then we get $AS = S\Lambda$ which implies $S^{-1}AS = \Lambda$.

There are some remarks about Λ

- **Remark 1** if the eigenvalues $\lambda_1, \dots, \lambda_n$ are all different. Then it is automatic that the eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are independent.⁸ And any matrix with no repeated eigenvalues can be diagonalized.
- **Remark 2** Some matrix with too few eigenvalues will make them undiagonalizable.
- **Remark 3** The eigenvectors are not unique, in other words, one can multiply some *nonzero* constant.

⁷without n independent eigenvectors, we can't diagonalize.

⁸The reason is that, if we try to express $\mathbf{0}$, for example, and consider only two eigenvalues λ_1, λ_2 and also two eigenvectors $\mathbf{x}_1, \mathbf{x}_2$. And we have $c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 = \mathbf{0}$, **multiplied by A** and **multiplied by λ_2** , we'll get $c_1 \lambda_1 \mathbf{x}_1 + c_2 \lambda_2 \mathbf{x}_2 = \mathbf{0}$ and $c_1 \lambda_2 \mathbf{x}_1 + c_2 \lambda_2 \mathbf{x}_2 = \mathbf{0}$, and then we get $c_1 = 0, c_2 = 0$, which means there is only one way to express $\mathbf{0}$.

6 RREF

RREF stands for *Reduced Row Echelon Form*. Before introducing it, let's talk about REF, row Echelon form. Here are the conditions,

- Each *nonzero* row lies above every *zero* row.
- Leading entries are Echelon form, shown below,

$$\begin{bmatrix} 1 & 0 & 0 & 6 & 0 & 3 \\ 0 & 0 & 1 & 5 & 0 & 7 \\ 0 & 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The leading entries are **the first nonzero number** in each row. Echelon form is that the leading entries in each nonzero row form like a stair going down from left to right.

RREF inherits the properties of REF with additional condition, that is, the columns containing leading entries are columns standard vector⁹. For example,

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

As we can see, the 1st, 3rd, and 4th columns with leading entries are standard vectors. It can be proved that RREF is unique.

⁹standard vector can be seen as unit vector.