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,

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

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 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 origined 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 comple number). And we call λ A's **eigenvalue** and **x** A's **eigenvector**.

In **Dirac Notaion**, a vector $\mathbf{x}$ can been seen as a **ket**[2] and the matirx A can been seen as an **operator**. There Dirac notaion is as shown below:

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

[1] with n rows and m columns.

[2] However, 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.

3 The Way to Find Eigenvalues and Eigenvectors

As we may see in the following secitons, 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 $det(A - \lambda I) = 0$

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

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*[3]. And since the definition of eigenvector is as below:

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

It is clearly[4] that

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

So we can conclude that **eigenvectors** are in the **nullspace**[5] of $A - \lambda I$, that is, we'll have the following equation for the **example**:

$$(A - I)\mathbf{x_1} = \mathbf{0}$$

[3]The definition of singular matrix is that the determination is 0, that is, not invertible

[4]one can check this by writing out the system of equations

[5]Nullspace 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 adn multiply without leaving the nullspace, so it is a subsapce.

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 introdution to eigenvalue and eigenvector, one should know that *special matrix will have special eigenvalue and eigenvector whose patterns and properties are worth studying*.

4 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 independent[6] eigenvectors $\mathbf{x_1}, ..., \mathbf{x_n}$.
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:

$$AS = A\begin{pmatrix} \mathbf{x_1} & ... & \mathbf{x_n} \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{x_1} & ... & \lambda_n\mathbf{x_n} \end{pmatrix}$$

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.

[6] without n independent eigenvectors, we can't diagonalize.

$$\begin{pmatrix}\lambda_1\mathbf{x_1} & ... & \lambda_n\mathbf{x_n}\end{pmatrix} = \begin{pmatrix}\mathbf{x_1} & ... & \mathbf{x_n}\end{pmatrix}\begin{pmatrix}\lambda_1 & 0 & ... & 0 & ... & \\ 0 & \lambda_2 & ... & 0 & ... & \\ & & . & & & \\ & & & . & & \\ & & & & . & \\ & & & & & \lambda_n\end{pmatrix} = S\Lambda$$

then we get $AS = S\Lambda$ which implies $S^{-1}AS = \Lambda$.
There are some remarks about Λ

- **Remark 1** if the eigenvalues $\lambda_1, ..., \lambda_n$ are all different. Then it is automatic that the eigenvectors $\mathbf{x_1}, ..., \mathbf{x_n}$ are independent.[7] 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.

[7]The reason is that, if we try to express 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 **0**.