Linear algebra

PSAAD

1 Eigenvector & Eigenvalue

Definition 1.1. *Given a vector space V over $\mathbb{C}$ and a linear transformation (matrix) $T : V \to V$. Let $|v\rangle \neq 0 \in V$, if there exists a scalar $\lambda \in \mathbb{C}$ s.t $T|v\rangle = \lambda |v\rangle$, then we call $|v\rangle$ is an eigenvector of T and λ is the corresponding eigenvalue.*

In practice, we can find the eigenvalues and eigenvectors of a given matrix A by the following way:

First, since $A|v\rangle = \lambda|v\rangle$, you can get $(A - \lambda I)|v\rangle = 0$. By the assumption $|v\rangle \neq 0$, we obtain $det(A - \lambda I) = 0$. You can solve the possible λ by $det(A - \lambda I) = 0$. Of course, you need to consider all the possible λ since there can be lots of eigenvalues. (There should be n λ if the matrix is $n \times n$, including repeated roots)

Second, you need to insert possible λ into $(A - \lambda I)|v\rangle = 0$ one after another to find the eigenvector. (There should be many possible eigenvectors since if $|v\rangle$ is an eigenvector of λ, then so does $a|v\rangle$. In most cases, you only need to find the eigenvectors which are linearly-indep.)

If you still don't know how to do, we're glad to see you searching relative problems online.

2 Diagonalization

Given a matrix $A \in M_n(\mathbb{C})$, if there exists n eigenvectors $|v_i\rangle \in V$, $i = 1, 2, \ldots, n$, with corresponding λ which forms a basis of V, then we can construct a matrix Q:

$$Q = \begin{pmatrix} |v_1\rangle & |v_2\rangle & |v_3\rangle & \ldots & |v_n\rangle \end{pmatrix}$$

and a diagonal matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots \\ 0 & \lambda_2 & \ldots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

then we can calculate:

$$AQ = \begin{pmatrix} \lambda_1 \left|v_1\right\rangle & \lambda_2 \left|v_2\right\rangle & \lambda_3 \left|v_3\right\rangle & \dots & \lambda_n \left|v_n\right\rangle \end{pmatrix} = Q\Lambda$$

Since $\left|v_1\right\rangle, \dots, \left|v_n\right\rangle$ is a basis of V, Q is invertible. Therefore, we have a conclusion $Q^{-1}AQ = \Lambda$.

This process is called the diagonalization. Diagonalization is very useful. For example, if you want to calculate A^{100}, you can replace A from $Q\Lambda Q^{-1}$. And all you have to do is to calculate $Q\Lambda^{100}Q^{-1}$, which is very easy.

To be careful: Although in most physical problems, we can diagonalize the target matrix. But, it is not always true because you might not be able to find the enough linearly-independent vectors to form a basis. For some matrices, you can only block-diagonalize them by Jordon Matrix. In the following section, we'll introduce which kinds of matrices can always be diagonized.

3 Spectral Theorem

3.1 Some definitions

Definition 3.1. ***(Invariant Subspace)*** *Let $\mathcal{V}$ be a vector spcae and $A : \mathcal{V} \to \mathcal{V}$ a linear map. If $\chi \subseteq \mathcal{V}$ and $A(\chi) \subseteq \chi$, then we say χ is an invariant subspace of A.*

Definition 3.2. *For a matrix A, if there is a polynomial $m_A(t)$, such that $m_A(A) = 0$, then $m_A(t)$ is called an* ***annihilating polynomial****. The annihilating polynomial with minimal degree is called* ***minimal polynomial****.*

Definition 3.3. *An inner product structure is a binary operator $\langle v_1, v_2\rangle : V \times V \longrightarrow \mathbb{C}$, which would satisfy the following properties:*

1. *Conjugate Symmetry: $\forall\, v_1,\, v_2 \in V,\ \langle v_1, v_2\rangle = \langle v_2, v_1\rangle^*$*
2. *Linearity: $\forall\, v_1,\, v_2,\, v_3 \in V,\ a \in \mathbb{C},\ \langle av_1 + v_3, v_2\rangle = a\langle v_1, v_2\rangle + \langle v_3, v_2\rangle$*
3. *Positive definiteness: $\forall v \neq 0 \in V,\ \langle v, v\rangle > 0$*

Remark: You can check the inner product you learn in high school is consistent with the definition 3.3, which is called the **standard complex inner product**. But there are still lots of inner product structure that you don't know as long as it satisfies definition 3.3. If we don't emphasize what inner product we use, we are using the standard inner product.

Definition 3.4. *The conjugate transpose (or adjoint) of a matrix A is denoted by $A^\dagger$ and defined by $A^\dagger := (A^t)^*$.*

Proposition 3.5. *Given a standard complex inner product structure $\langle \cdot, \cdot \rangle$, a matrix A and two vectors v_1, v_2, we have $\langle Av_1, v_2 \rangle = \langle v_1, A^\dagger v_2 \rangle$.*

Definition 3.6. *In this definition, we restrict our consideration from $\mathbb{C}$ to $\mathbb{R}$. For a matrix O, we call it an orthogonal matrix if O preseves the inner product value. In other words, $\forall v_1, v_2 \in V$, $\langle Ov_1, Ov_2 \rangle = \langle v_1, v_2 \rangle$. Therefore, you can conclude that $OO^t = O^tO = I$*

Definition 3.7. *For a matrix Q, we call it a normal matrix if $QQ^\dagger = Q^\dagger Q$.*

Definition 3.8. *For a matrix U, we call it an unitary matrix if $UU^\dagger = U^\dagger U = I$.*

Definition 3.9. *For a matrix U, we call it a hermitian matrix if $U = U^\dagger$.*

3.2 Useful Statements

Theorem 3.10. ***(Schur's Theorem)*** *For any complex square matrix S, there exists an unitary matrix U, such that $T = USU^\dagger$ where T is a upper triangluation matrix.*

Proposition 3.11. *Let A be Hermitian. If χ is an invariant subspace of A, then $\chi^\perp$ (complement space, in case you don't know what it is.) is also an invariant subspace of A.*

Proof. Let $x \in \chi$ and $y \in \chi^\perp$. Since χ is an invariant subspace, $Ax \in \chi$. We have

$$\langle y|Ax \rangle = 0 \Longleftrightarrow \langle Ay|x \rangle = 0 \Longleftrightarrow Ay \in \chi^\perp$$

□

Theorem 3.12. *A matrix S is* ***diagonalizable with a unitary matrix*** *iff S is a* ***normal matrix****.*

Proof. ($\Longrightarrow$) According to **Schur's theorem**, any square matrix can be triangulated by unitary transformation : $T = USU^\dagger$ where T is an upper triangulation matrix. If S is normal, then

$$\begin{aligned} TT^\dagger &= (U^\dagger SU)(U^\dagger SU)^\dagger \\ &= U^\dagger SS^\dagger U \\ &= U^\dagger S^\dagger SU \\ &= T^\dagger T \end{aligned}$$

Thus, T is normal as well. We can write out $TT^\dagger = T^\dagger T$ explicitly:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \cdots & \end{pmatrix} \begin{pmatrix} a_{11}^* & & \\ a_{12}^* & a_{22}^* & \\ a_{13}^* & a_{23}^* & a_{33}^* \\ \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} a_{11}^* & & \\ a_{12}^* & a_{22}^* & \\ a_{13}^* & a_{23}^* & a_{33}^* \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \cdots & \end{pmatrix}$$

For the 11 entry, we have

$$|a_{11}|^2 + |a_{12}|^2 + |a_{13}|^2 + \cdots = |a_{11}|^2$$

Manifestly, it's only possible if $a_{12} = a_{13} = \cdots = 0$. Similarly, for any diagonal entries, we find T is, in fact, diagonized matrix. As a result, S is diagonizable with an unitary matrix.

($\Longleftarrow$)Let $\hat{S}$ be the unitarily diagonized matrix from S. We have $\hat{S} = USU^\dagger$. Evidentally,

$$SS^\dagger = U^\dagger \hat{S}\hat{S}^\dagger U = U^\dagger \hat{S}\hat{S}^\dagger U = SS^\dagger$$

$\square$

Proposition 3.13. *If S is a normal matrix:*

- *If $S|\lambda\rangle = \lambda|\lambda\rangle$, then $S^\dagger = \lambda^*|\lambda\rangle$*
- *For different eigenvalues, the corresponding eigenvectors are orthogonal,i.e. $\langle\lambda|\lambda'\rangle = 0$ if $\lambda \neq \lambda'$.*

Since if a matrix is hermitian, it's normal, any hermitian matrix can be diagonized by an unitary matrix. Following we give an alternative proof for this fact for hermitian operator exclusively. We first give a lemma.

Lemma 3.14. *Let $\mathcal{L} : \mathbb{C}^n \to \mathbb{C}^n$ be a linear map. If $\chi \subseteq \mathbb{C}^n$ is an invariant subspace of $\mathcal{L}$, then there is a nonzero vector $|x\rangle \in \chi$ such that $\mathcal{L}|x\rangle = \lambda|x\rangle$, $\lambda \in \mathbb{C}$.*

Proof. Let $dim\chi = r \leq n$ and $|y\rangle \in \chi$ where $|y\rangle \neq 0$. Note that $\{|y\rangle, \mathcal{L}|y\rangle, \cdots, \mathcal{L}^r|y\rangle\}$ is linear dependent. Thus, there are non all zero $c_i, i = 0, \cdots, r$ such that :

$$(c_0 + c_1\mathcal{L} + \cdots + c_r\mathcal{L}^r)|y\rangle = 0$$

We can factorize the equation:

$$c_r(\mathcal{L} - \mu_1)\cdots(\mathcal{L} - \mu_r)|y\rangle = 0$$

As a result, there exist at least a $|x\rangle = \Pi_{i\in I}(\mathcal{L} - \mu_i)|y\rangle \neq 0$ for some subset I of $1, \cdots, r$, such that

$$\begin{aligned} (\mathcal{L} - \mu_k)|x\rangle &= 0 \\ \Longrightarrow \mathcal{L}|x\rangle &= \mu_k|x\rangle \end{aligned}$$

for some k. $\square$

Theorem 3.15. *Let $\mathcal{L}$ be a* ***Hermitian operator*** *acting on a n dimension vector space. Then $\mathcal{L}$ has a completely orthonormal eigenvectors basis.*

Proof. Assume $\mathcal{L}$ has up to $k(0 \leq k < n)$ orthonormal eigenvectors $S = \{|x_1\rangle \cdots |x_k\rangle\}$. Let $\chi = span\{S\}$. χ is an invariant subspace of $\mathcal{L}$. Since $\mathcal{L} = \mathcal{L}^\dagger$, $\chi^\perp$ is also an invariant subspace of $\mathcal{L}$. But, $k < n$; hence, $\chi^\perp \neq 0$. By lemma, there is a nonzero $y \in \chi^\perp$(can be normalized) such that $\mathcal{L}|y\rangle = \lambda|y\rangle$. Thus, $\{|y\rangle, |x_1\rangle, \cdots, |x_k\rangle\}$ is an orthnormal eigenvector basis, leading to contradiction. Hence, $k = n$. $\square$