Spectral theorem

Jiman Hwang

This document is for self-study only.
Source: [Nie10]

1 Quick definitions

Conjugate transpose:

$$M^\dagger = \overline{M}^T$$

Dirac bra-ket notations:

$$|v\rangle = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},\ \langle v| = |v\rangle^\dagger$$

and inner product:

$$\langle v \,|\, w\rangle$$

2 Statement

Any normal operator M on a vector space V is unitarily diagonalizable if and only if M is normal. That is, given a $n \times n$ matrix M,

$$M^\dagger M = MM^\dagger \quad \Leftrightarrow \quad \exists U \exists D \quad M = UDU^\dagger \tag{1}$$

where U is some unitary matrix($U^\dagger = U^{-1}$) whose columns consist of eigenvectors of M, and $D = \mathbf{diag}(\lambda_1, \cdots, \lambda_n)$ where $\lambda_1, \cdots, \lambda_n$ are corresponding eigenvalues of M.

3 Derivation

3.1 Normal matrix $\Leftarrow$ unitarily diagonalizable

If

$$M = UDU^\dagger$$

then,

$$\begin{aligned}
M^\dagger M =& \left(UDU^\dagger\right)^\dagger \left(UDU^\dagger\right) \\
=& UD^\dagger U^\dagger UDU^\dagger \\
=& UD^\dagger DU^\dagger \\
=& UDD^\dagger U^\dagger \quad \text{(diagonal)} \\
=& UDU^\dagger UD^\dagger U^\dagger \quad \text{(diagonal)} \\
=& UDU^\dagger \left(UDU^\dagger\right)^\dagger \\
=& MM^\dagger
\end{aligned}$$

3.2 Normal matrix $\Rightarrow$ unitarily diagonalizable

3.2.1 Some terminologies and theorems

Use induction for dimension of V. $\dim V = 1$ is trivial. Now assume (1) is held for $\dim V < n$, and we will show (1) is held for $\dim V = n$. Let

- λ is a eigenvalue of normal matrix M
- P is the projector onto the λ eigenspace, W, whose dimension is k
- Q is the projector onto $W^\perp$
- $\{|v_1\rangle, \cdots, |v_k\rangle\}$ is an orthonormal basis of W
- $\{|v_{k+1}\rangle, \cdots, |v_n\rangle\}$ is an orthonormal basis of $W^\perp$

We can express P and Q as follow.

$$P = \sum_{i=1}^{k} |v_i\rangle \langle v_i|, \; Q = \sum_{i=k+1}^{n} |v_i\rangle \langle v_i|$$

Note that

$$P + Q = I \tag{2}$$

Because

$$\forall |x\rangle \in V \quad \exists \{a_1, \cdots, a_n\} \quad |x\rangle = \sum_{i=1}^{n} a_i |v_i\rangle$$

Calculating $(P+Q)|x\rangle$,

$$(P+Q)|x\rangle = \sum_{i=1}^{n} |v_i\rangle \langle v_i|x\rangle = \sum_{i=1}^{n} a_i |v_i\rangle = |x\rangle$$

Since $|x\rangle$ is arbitrary, (2) is held. Also note

$$P^2 = P, \; P^\dagger = P, \; Q^2 = Q, \; Q^\dagger = Q \tag{3}$$

Proofs for those are skipped.

3.2.2 The proof

Using (2),

$$M = (P+Q)M(P+Q) = PMP + PMQ + QMP + QMQ \tag{4}$$

Let's inspect each term one by one. For PMP,

$$PMP = P \cdot \lambda P = \lambda P^2 = \lambda P \tag{5}$$

For PMQ,

$$\forall |x\rangle \in V \quad MM^\dagger P|x\rangle = M^\dagger MP|x\rangle = \lambda M^\dagger P|x\rangle$$

Thus, $M^\dagger P|x\rangle$ is in W. Also, projecting it onto $W^\perp$ produces zero vector. That is,

$$QM^\dagger P|x\rangle = \mathbf{0}$$

Since this is held for any $|x\rangle$,

$$QM^\dagger P = \mathbf{0}$$

Applying conjugate transpose and using (3),

$$QM^\dagger P = \mathbf{0} \quad \Rightarrow \quad P^\dagger MQ^\dagger = \mathbf{0} \quad \Rightarrow \quad PMQ = \mathbf{0} \tag{6}$$

For QMP,

$$QMP = Q \cdot \lambda P = \lambda QP = \mathbf{0} \tag{7}$$

Using (4), (5), (6), and (7),

$$M = \lambda P + QMQ \tag{8}$$

We show that λP and QMQ are diagonal with respect to some orthonormal bases for W and $W^\perp$ respectively. Then, assembling those decomposed matrices ends the proof. For λP,

$$\begin{aligned}
\lambda P &= \sum_{i=1}^{k} \lambda |v_i\rangle \langle v_i| \\
&= \sum_{i=1}^{k} |v_i\rangle \lambda I \langle v_i| \\
&= \begin{bmatrix} |v_1\rangle & \cdots & |v_k\rangle \end{bmatrix} \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix} \begin{bmatrix} \langle v_1| \\ \vdots \\ \langle v_k| \end{bmatrix}
\end{aligned}$$

Let

$$U_1 = \begin{bmatrix} |v_1\rangle & \cdots & |v_k\rangle \end{bmatrix},\ D_1 = \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix} \quad (k \times k)$$

Then,

$$\lambda P = U_1 D_1 U_1^\dagger \tag{9}$$

For QMQ, we state QMQ is normal.

$$\begin{aligned} QMQ &= (I-P)MQ = MQ \quad \text{by (6)} \\ QMQ &= QM(I-P) = QM \quad \text{by (7)} \end{aligned}$$

Using those,

$$\begin{aligned} (QMQ)^\dagger (QMQ) =& (MQ)^\dagger (MQ) \\ =& Q^\dagger M^\dagger MQ \\ =& Q^\dagger MM^\dagger Q \quad M \text{ is normal.} \\ =& QMM^\dagger Q^\dagger \quad \text{by (3)} \\ =& QM(QM)^\dagger \\ =& (QMQ)(QMQ)^\dagger \end{aligned}$$

Thus, QMQ is normal. Also, since $QMQ = QM$ is on $W^\perp$ and $\dim W \geq 1$,

$$\operatorname{rank}(QMQ) < n$$

By induction, it is true that every normal matrix on a vector space whose dimension is less than n is unitarily diagonalizable. Thus,

$$\exists U_2 \exists U_2' \exists D_2 \quad QMQ = \begin{bmatrix} U_2 & U_2' \end{bmatrix} \begin{bmatrix} D_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} U_2^\dagger \\ (U_2')^\dagger \end{bmatrix}$$

where U_2 is a $n \times (n-k)$ matrix whose column vectors are some orthonormal basis of $W^\perp$, D_2 is a $(n-k) \times (n-k)$ diagonal matrix U_2' is a $n \times k$ matrix such that $\begin{bmatrix} U_2 & U_2' \end{bmatrix}$ is unitary. The bottom-right element is $\mathbf{0}$ in the 2nd matrix because when QMQ is transformed with respect to column vectors in U_2, there's no portion parallel to W. Calculating the above expression,

$$QMQ = U_2 D_2 U_2^\dagger \tag{10}$$

Assembling (8), (9), and (10),

$$\begin{aligned} M =& U_1 D_1 U_1^\dagger + U_2 D_2 U_2^\dagger \\ =& \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} D_1 & \mathbf{0} \\ \mathbf{0} & D_2 \end{bmatrix} \begin{bmatrix} U_1^\dagger \\ U_2^\dagger \end{bmatrix} \\ =& UDU^\dagger \end{aligned}$$

Note that column vectors in U are orthonormal, hence U is unitary, and D is diagonal. In addition, if

$$U = \begin{bmatrix} |w_1\rangle & \cdots & |w_n\rangle \end{bmatrix}, \; D = \mathbf{diag}(\lambda_1, \cdots, \lambda_n)$$

then,

$$
\begin{aligned}
& M = UDU^{\dagger} \\
\Rightarrow\ & MU = UD \\
\Rightarrow\ & \begin{bmatrix} M\left|w_1\right\rangle & \cdots & M\left|w_n\right\rangle \end{bmatrix} = \begin{bmatrix} \lambda_1\left|w_1\right\rangle & \cdots & \lambda_n\left|w_n\right\rangle \end{bmatrix}
\end{aligned}
$$

Thus, λ_i is an eigenvalue of M and $|w_i\rangle$ is the corresponding eigenvector. This finishes the proof.

References

[Nie10] Nielsen, Michael A., and Isaac L. Chuang. *Quantum Computation and Quantum Information*, page 72. Cambridge University Press, 1th edition, 2010.