

Linear Algebra Class on 8 June

Seanie Lee

15 June 2019

6.6 Orthogonal Projection and Spectral Theorem

Definition. $A \in \mathfrak{M}_{n \times n}(\mathbb{R})$, A is orthogonal if $A^t A = AA^t = \mathbf{I}_n$
 $A \in \mathfrak{M}_{n \times n}(\mathbb{C})$, A is unitary if $A^* A = AA^* = \mathbf{I}_n$

Remark.

$$\begin{aligned} A = [\mathcal{C}_1 \cdots \mathcal{C}_n] \text{ is unitary} &\iff [A^* A]_{ij} = \sum_{k=1}^n A_{ik}^* A_{kj} \\ &= \sum_{k=1}^n \overline{A_{ki}}^t A_{kj} \\ &= \mathcal{C}_i^* \mathcal{C}_j \\ &= \mathcal{C}_j \cdot \mathcal{C}_i \\ &= \delta_{ji} \end{aligned}$$

Thus it is equivalent to $\{\mathcal{C}_1, \dots, \mathcal{C}_n\}$ is an orthonormal basis for $\mathbb{C}^n(\mathbb{R}^n)$

Definition.

A is unitarily equivalent to $B \iff B = U^{-1} A U$ for some unitary matrix $U (\cdot : U^* U = \mathbf{I}_n \implies U^* = U^{-1})$

A is unitarily diagonalizable $\iff A$ is unitarily equivalent to diagonal matrix $D, A = U^{-1} D U$

Theorem. β : an orthonormal basis consisting of eigenvectors of $\mathbf{T} \iff \mathbf{T}$ is unitarily diagonalizable

Proof. \implies since \mathbf{T} is unitarily diagonalizable, $[\mathbf{T}]_\gamma = Q^{-1} [\mathbf{T}]_\beta Q$. $Q := [\mathbf{x}_1 \cdots \mathbf{x}_n]$, $[\mathbf{T}]_\beta := \text{diag}(\lambda_1, \dots, \lambda_n)$. The $\beta = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is the set of orthonormal eigenvectors.

$\therefore Q$ is unitary matrix.

\Leftarrow Since \mathbf{T} is unitarily diagonalizable, $[\mathbf{T}]_\gamma = U^{-1} D U$ for some unitary matrix U and some diagonal matrix D . Thus $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ where λ_i is eigenvalue and column vectors of U are orthonormal basis.

\therefore There exists an orthonormal basis consisting of eigenvectors of \mathbf{T} \square

Definition (Projection \mathbf{T}). $\mathbf{W}_1, \mathbf{W}_2$ are subspaces of \mathbf{V} , $\mathbf{V} = \mathbf{W}_1 \oplus \mathbf{W}_2$, and $\mathbf{T} : \mathbf{V} \longrightarrow \mathbf{V}$ is projection on \mathbf{W}_1 along \mathbf{W}_2 if $\mathbf{T}\mathbf{x} = \mathbf{x}_1$ where $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \in \mathbf{W}_1 \oplus \mathbf{W}_2$.

Note. $\text{im}\mathbf{T} = \mathbf{W}_1 = \{\mathbf{x} \in \mathbf{V} \mid \mathbf{T}\mathbf{x} = \mathbf{x}\}$ and $\ker \mathbf{T} = \mathbf{W}_2$. $\therefore \mathbf{V} = \text{im}\mathbf{T} \oplus \ker \mathbf{T}$. Thus there is no ambiguity if we refer to \mathbf{T} as a "projection on \mathbf{W}_1 " or simply as a projection.

Note. Because $\mathbf{V} = \mathbf{W}_1 \oplus \mathbf{W}_2 = \mathbf{W}_1 \oplus \mathbf{W}_3$ does not imply that $\mathbf{W}_2 = \mathbf{W}_3$, \mathbf{W}_1 does not uniquely determine \mathbf{T} . However, for an orthogonal projection \mathbf{T} , \mathbf{T} is uniquely determined by its range.

Theorem. T is projection if and only if $\mathbf{T}^2 = \mathbf{T}$

Proof. $\mathbf{x} := \mathbf{x}_1 + \mathbf{x}_2 \in \mathbf{W}_1 \oplus \mathbf{W}_2$, thus $\mathbf{T}\mathbf{x} = \mathbf{x}_1$. Since $\mathbf{T}^2\mathbf{x} = \mathbf{T}\mathbf{x}_1 = \mathbf{x}_1 = \mathbf{T}\mathbf{x}$, $\mathbf{T}^2\mathbf{x} = \mathbf{T}\mathbf{x} \ \forall \mathbf{x} \in \mathbf{V}$.
 $\therefore \mathbf{T}^2 = \mathbf{T}$

\Leftarrow Suppose $\mathbf{T}^2 = \mathbf{T}$. $\mathbf{x} = \mathbf{T}\mathbf{x} + (\mathbf{x} - \mathbf{T}\mathbf{x})$. Then $\mathbf{T}\mathbf{x} \in \text{im}\mathbf{T}$. We want to show that $\mathbf{x} - \mathbf{T}\mathbf{x} \in \ker \mathbf{T}$.

$$\begin{aligned} \mathbf{T}(\mathbf{x} - \mathbf{T}\mathbf{x}) &= \mathbf{T}\mathbf{x} - \mathbf{T}^2\mathbf{x} \\ &= \mathbf{T}\mathbf{x} - \mathbf{T}\mathbf{x} \\ &= \mathbf{0} \\ \therefore \mathbf{x} - \mathbf{T}\mathbf{x} &\in \ker \mathbf{T} \end{aligned}$$

Now we want to show that $\text{im}\mathbf{T} \cap \ker \mathbf{T} = \mathbf{0}$. Let $\mathbf{x} \in \text{im}\mathbf{T} \cap \ker \mathbf{T}$ be given. Since $\mathbf{x} \in \ker \mathbf{T}$, $\mathbf{T}\mathbf{x} = \mathbf{0}$. Since $\mathbf{x} \in \text{im}\mathbf{T}$, $\exists \mathbf{x}_0 \in \mathbf{V}$ such that $\mathbf{x} = \mathbf{T}\mathbf{x}_0 = \mathbf{T}^2\mathbf{x}_0 = \mathbf{T}(\mathbf{T}\mathbf{x}_0) = \mathbf{T}\mathbf{x} = \mathbf{0}$. Thus $\text{im}\mathbf{T} \cap \ker \mathbf{T} = \mathbf{0}$, i.e., $\mathbf{V} = \text{im}\mathbf{T} \oplus \ker \mathbf{T}$. Put $\mathbf{W}_1 := \text{im}\mathbf{T}$, $\mathbf{W}_2 := \ker \mathbf{T}$. $\therefore \mathbf{T}$ is projection on \mathbf{W}_1 \square

Definition. \mathbf{T} is orthogonal projection of \mathbf{V} on \mathbf{W} if and only if $\text{im}\mathbf{T} = \mathbf{W}$ and $\ker \mathbf{T} = \text{im}\mathbf{T}^\perp$, $\ker \mathbf{T}^\perp = \text{im}\mathbf{T}$

Theorem. \mathbf{T} is orthogonal projection on $\mathbf{W} \iff \exists \mathbf{T}^*$ and $\mathbf{T}^2 = \mathbf{T} = \mathbf{T}^*$

Proof. $\implies \mathbf{T}$ is projection $\iff \mathbf{T}^2 = \mathbf{T}$. It is enough to show that $\exists \mathbf{T}^*$ such that $\mathbf{T}^* = \mathbf{T}$. $\mathbf{V} = \text{im}\mathbf{T} \oplus \ker \mathbf{T}$ and $\text{im}\mathbf{T}^\perp = \ker \mathbf{T}$, $\ker \mathbf{T}^\perp = \text{im}\mathbf{T}$. Let $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ with $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$ for $\mathbf{x}_1, \mathbf{y}_1 \in \text{im}\mathbf{T}$ for $\mathbf{x}_2, \mathbf{y}_2 \in \ker \mathbf{T}$

$$\begin{aligned} \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle &= \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{T}(\mathbf{y}_1 + \mathbf{y}_2) \rangle \\ &= \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{T}\mathbf{y}_1 \rangle \\ (\because \mathbf{y}_2 &\in \ker \mathbf{T}) \\ &= \langle \mathbf{x}_1, \mathbf{T}\mathbf{y}_1 \rangle + \langle \mathbf{x}_2, \mathbf{T}\mathbf{y}_1 \rangle \\ &= \langle \mathbf{x}_1, \mathbf{y}_1 \rangle + \langle \mathbf{x}_2, \mathbf{y}_1 \rangle \\ (\because \mathbf{y}_1 &\in \text{im}\mathbf{T}, \mathbf{T}^2 = \mathbf{T}, \mathbf{y}_1 = \mathbf{T}\mathbf{y}' = \mathbf{T}^2\mathbf{y}' = \mathbf{T}\mathbf{y}_1) \end{aligned}$$

Since $\ker \mathbf{T} = \text{im}\mathbf{T}^\perp$, $\langle \mathbf{x}_2, \mathbf{y}_1 \rangle = 0$. i.e., $\langle \mathbf{x}_1, \mathbf{T}\mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y}_1 \rangle$

$$\begin{aligned} \langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{T}\mathbf{x}_1 + \mathbf{T}\mathbf{x}_2, \mathbf{y}_1 + \mathbf{y}_2 \rangle \\ &= \langle \mathbf{T}\mathbf{x}_1, \mathbf{y}_1 + \mathbf{y}_2 \rangle \\ &= \langle \mathbf{x}_1, \mathbf{y}_1 \rangle + \langle \mathbf{x}_1, \mathbf{y}_2 \rangle \\ &= \langle \mathbf{x}_1, \mathbf{y}_1 \rangle \\ \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle &= \langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

Put $\mathbf{T}^* := \mathbf{T}$. $\therefore \exists \mathbf{T}^*$ such that $\mathbf{T}^* = \mathbf{T}$

\Leftarrow since $\mathbf{T}^2 = \mathbf{T}$, \mathbf{T} is projection. We want to show that $\text{im}\mathbf{T} = \ker \mathbf{T}^\perp$ and $\ker \mathbf{T} = \text{im}\mathbf{T}^\perp$. Let

$\mathbf{x} \in \text{im } \mathbf{T}, \mathbf{y} \in \ker \mathbf{T}$. Since $\mathbf{T}^2 = \mathbf{T}$ and $\mathbf{x} \in \text{im } \mathbf{T}$, $\mathbf{T}\mathbf{x} = \mathbf{T}^2\mathbf{x}_0 = \mathbf{T}\mathbf{x}_0 = \mathbf{x}$. Then $\mathbf{x} = \mathbf{T}\mathbf{x} = \mathbf{T}^*\mathbf{x}$

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle \\ &= \langle \mathbf{T}^*\mathbf{x}, \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{0} \rangle \\ &= 0 \\ \therefore \mathbf{x} &\in \ker \mathbf{T}^\perp\end{aligned}$$

Let $\mathbf{y} \in \ker \mathbf{T}^\perp$. We want to show that $\mathbf{T}\mathbf{y} = \mathbf{y}$

$$\begin{aligned}\|\mathbf{y} - \mathbf{T}\mathbf{y}\|^2 &= \langle \mathbf{y} - \mathbf{T}\mathbf{y}, \mathbf{y} - \mathbf{T}\mathbf{y} \rangle \\ &= \langle \mathbf{y}, \mathbf{y} - \mathbf{T}\mathbf{y} \rangle - \langle \mathbf{T}\mathbf{y}, \mathbf{y} - \mathbf{T}\mathbf{y} \rangle\end{aligned}$$

Since $\mathbf{y} - \mathbf{T}\mathbf{y} \in \ker \mathbf{T}$, $\langle \mathbf{y}, \mathbf{y} - \mathbf{T}\mathbf{y} \rangle = 0$.

$$\begin{aligned}\langle \mathbf{T}\mathbf{y}, \mathbf{y} - \mathbf{T}\mathbf{y} \rangle &= \langle \mathbf{y}, \mathbf{T}^*(\mathbf{y} - \mathbf{T}\mathbf{y}) \rangle \\ &= \langle \mathbf{y}, \mathbf{T}(\mathbf{y} - \mathbf{T}\mathbf{y}) \rangle \\ &= \langle \mathbf{y}, \mathbf{0} \rangle \\ &= 0 \\ \therefore \mathbf{y} &= \mathbf{T}\mathbf{y}\end{aligned}$$

Then we want to show that $\ker \mathbf{T} = \text{im } \mathbf{T}^\perp$. $\ker \mathbf{T} \subset (\ker \mathbf{T}^\perp)^\perp = \text{im } \mathbf{T}^\perp$.

Let $\mathbf{x} \in \text{im } \mathbf{T}^\perp$. $\forall \mathbf{y} \in \mathbf{V}$,

$$\begin{aligned}\langle \mathbf{T}, \mathbf{y} \rangle &= \langle \mathbf{x}, \mathbf{T}^*\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle \\ &= 0 \\ \therefore \mathbf{T}\mathbf{x} &= \mathbf{0}\end{aligned}$$

Thus $\mathbf{x} \in \ker \mathbf{T}$, $\therefore \ker \mathbf{T} = \text{im } \mathbf{T}^\perp$ □

Note. \mathbf{T} is orthogonal projection on \mathbf{W} . $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthonormal basis of \mathbf{V} such that $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is an orthonormal basis for $\mathbf{W} \leq \mathbf{V}$. Then $[\mathbf{T}]_\beta = \text{diag}(1, \dots, 1, 0, \dots, 0)$.

$$\mathbf{T}\mathbf{v}_i = \begin{cases} \mathbf{v}_i & (i \leq k) \\ \mathbf{0} & (i > k) \end{cases}$$

Theorem (Spectral Theorem). \mathbf{V} is finite-dimensional vector space over \mathbf{F} and \mathbf{T} is linear operator on \mathbf{V} with distinct eigenvalue $1, \dots, \lambda_k$. \mathbf{T} is normal if $\mathbf{F} = \mathbb{C}$ or \mathbf{T} is self-adjoint if $\mathbf{F} = \mathbb{R}$. $\mathbf{W}_i = \mathbf{E}_{\lambda_i}$ for \mathbf{T}_i is orthogonal projection on \mathbf{W}_i .

(a). $\mathbf{V} = \mathbf{W}_1 + \bigoplus + \mathbf{W}_k$ (eigenspace decomposition)

(b). $\mathbf{W}'_i := \bigoplus_{j \neq i} \mathbf{W}_j \implies \mathbf{W}'_i = \mathbf{W}_i^\perp$

(c). $\mathbf{T}_i \circ \mathbf{T}_j = \delta_{ij} \mathbf{T}_i$

(d). $\mathbf{I} = \mathbf{T}_1 \cdots + \mathbf{T}_k$

(e). $\mathbf{T} = \lambda_1 \mathbf{T}_1 + \cdots + \lambda_k \mathbf{T}_k$ (Spectral decomposition)

(b). *Proof.* For $\mathbf{x} \in \mathbf{W}_i, \mathbf{y} \in \mathbf{W}_j, \langle \mathbf{x}, \mathbf{y} \rangle = 0$. Since \mathbf{T} is normal or self-adjoint, two eigenvectors corresponding each to two different eigenvalues are orthogonal. $\mathbf{W}'_i \subset \mathbf{W}_i^\perp$

$$\begin{aligned} \dim \mathbf{W}'_i &= \sum_{i \neq j} \dim \mathbf{W}_j \\ &= \dim \mathbf{V} - \dim \mathbf{W}_i \\ &= \dim \mathbf{W}_i^\perp \\ \mathbf{W}'_i &= \mathbf{W}_i^\perp \end{aligned}$$

□

(c). *Proof.* $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_k$ where $\mathbf{x}_i \in \mathbf{W}_i$. $\mathbf{T}_i \mathbf{T}_j(\mathbf{x}) = \mathbf{T}_i \mathbf{x}_j$

$$\begin{cases} i = j & \mathbf{T}_i \mathbf{x}_j = \mathbf{x}_i = \mathbf{T}_i \mathbf{x} \\ i \neq j & \mathbf{T}_i \mathbf{x}_j = \mathbf{0} \end{cases}$$

$\therefore \mathbf{T}_i \mathbf{T}_j = \delta_{ij} \mathbf{T}_i$

□

(d). *Proof.* $\ker \mathbf{T}_i = \text{im} \mathbf{T}_i^\perp = \mathbf{W}_i^\perp = \mathbf{W}'_i$. $\mathbf{V} = \text{im} \mathbf{T}_i + \ker \mathbf{T}_i (= \mathbf{W}'_i)$
For all $\mathbf{x} \in \mathbf{V}$,

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_i + (\mathbf{x}_1 + \cdots + \mathbf{x}_{i-1} + \mathbf{x}_{i+1} + \cdots + \mathbf{x}_k) \\ &= \mathbf{T}_i \mathbf{x} + (\mathbf{T}_1 \mathbf{x} + \cdots + \mathbf{T}_{i-1} \mathbf{x} + \mathbf{T}_{i+1} \mathbf{x} + \cdots + \mathbf{T}_k \mathbf{x}) \\ &= (\mathbf{T}_1 + \cdots + \mathbf{T}_k) \mathbf{x} \\ \therefore \mathbf{I} &= \mathbf{T}_1 + \cdots + \mathbf{T}_k \end{aligned}$$

□

(e). *Proof.* $\forall \mathbf{x} \in \mathbf{V}, \mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_k$ for $\mathbf{x}_i \in \mathbf{W}_i$

$$\begin{aligned} \mathbf{T} \mathbf{x} &= \mathbf{T}(\mathbf{x}_1 + \cdots + \mathbf{x}_k) \\ &= \mathbf{T} \mathbf{x}_1 + \cdots + \mathbf{T} \mathbf{x}_k \\ &= \lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k \\ &= \lambda_1 \mathbf{T}_1 \mathbf{x} + \cdots + \lambda_k \mathbf{T}_k \mathbf{x} \\ &= (\lambda_1 \mathbf{T}_1 + \cdots + \lambda_k \mathbf{T}_k) \mathbf{x} \\ \therefore &= \lambda_1 \mathbf{T}_1 + \cdots + \lambda_k \mathbf{T}_k \end{aligned}$$

□

Note. If A is unitarily diagonalizable, $A = Q^{-1} A Q = Q^* A Q$. Thus getting inverse of Q is just \overline{Q}^t . It's much computationally cheaper than usual case.

Note. If \mathbf{T} is normal (self-adjoint) and β is the union of orthonormal bases of the \mathbf{W}'_i s and let $M_i =$

$$\dim \mathbf{W}_i, \text{ then } [\mathbf{T}]_\beta = \begin{pmatrix} \lambda_1 \mathbf{I}_{m_1} & & & 0 \\ & \lambda_2 \mathbf{I}_{m_2} & & \\ & & \ddots & \\ 0 & & & \lambda_k \mathbf{I}_{m_k} \end{pmatrix}$$

Corollary. \mathbf{T} is unitary $\iff \mathbf{T}$ is normal and $|\lambda_i| = 1$ for $i = 1, \dots, k$

Proof. \implies Suppose that \mathbf{T} is unitary. Since $\mathbf{T}^* \mathbf{T} = \mathbf{T} \mathbf{T}^*$, \mathbf{T} is normal.

$$\begin{aligned} \mathbf{T} \mathbf{T}^* \mathbf{v}_i &= \mathbf{T}(\bar{\lambda}_i \mathbf{v}_i) \\ &= |\lambda_i|^2 \mathbf{v}_i \\ &= \mathbf{v}_i \quad (\because \mathbf{T} \mathbf{T}^* = \mathbf{I}) \\ \therefore |\lambda_i|^2 &= 1 \end{aligned}$$

\Leftarrow Suppose that \mathbf{T} is normal and $|\lambda_i| = 1$. Since \mathbf{T}_i is orthogonal projection, $\mathbf{T}_i^2 = \mathbf{T}_i = \mathbf{T}_i^*$.

$$\begin{aligned} \mathbf{T}^* \mathbf{T} &= (\bar{\lambda}_1 \mathbf{T}_1 + \dots + \bar{\lambda}_k \mathbf{T}_k)(\lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k) \\ &= |\lambda_1|^2 \mathbf{T}_1^2 + \dots + |\lambda_k|^2 \mathbf{T}_k^2 \\ &= \mathbf{T}_1 + \dots + \mathbf{T}_k \\ &= \mathbf{I} \end{aligned}$$

$\therefore \mathbf{T}$ is unitary □

Corollary. *mathbf{T}* is normal and $\mathbf{F} = \mathbb{C}$. Then \mathbf{T} is self-adjoint if and only if $\forall \lambda \lambda \in \mathbb{R}$

Proof. \implies Suppose that \mathbf{T} is self-adjoint. For all i

$$\begin{aligned} \lambda_i \mathbf{x}_i &= \mathbf{T} \mathbf{x}_i \\ &= \mathbf{T}^* \mathbf{x}_i \\ &= \bar{\lambda}_i \mathbf{x}_i \\ \therefore \bar{\lambda}_i &= \lambda_i \\ \therefore \lambda_i &\in \mathbb{R} \end{aligned}$$

\Leftarrow

$$\begin{aligned} \mathbf{T} &= \lambda_1 \mathbf{T}_1 + \lambda_2 \mathbf{T}_2 + \dots + \lambda_k \mathbf{T}_k \\ \mathbf{T}^* &= \bar{\lambda}_1 \mathbf{T}_1 + \dots + \bar{\lambda}_k \mathbf{T}_k \\ &= \lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k \\ &= \mathbf{T} \end{aligned}$$

$\therefore \mathbf{T}$ is self-adjoint □

Definition (Singular Value Decomposition). \mathbf{V}, \mathbf{W} are finite-dimensional inner product space. $\mathbf{T} : \mathbf{V} \rightarrow \mathbf{W}$: linear transformation of rank r . Then $\exists \beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$: an orthonormal basis for \mathbf{V} , $\gamma = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$: an orthonormal basis of \mathbf{W} , and $\sigma_1 \geq \sigma_2 \leq \dots \leq \sigma_r > 0$ such that

$$\mathbf{T}\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i & (i \leq r) \\ \mathbf{0} & (i > r) \end{cases}$$

Conversely if the preceding conditions are satisfied, then

$$\begin{cases} \mathbf{T}^* \mathbf{T} \mathbf{v}_i = \sigma_i^2 \mathbf{v}_i & (i \leq r) \\ \mathbf{T}^* \mathbf{T} \mathbf{v}_i = \mathbf{0} & (i > r) \end{cases}$$

\therefore the scalars $\sigma_1, \dots, \sigma_r$ are uniquely determined by \mathbf{T}