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.

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

Thus it is equivalent to $\{C_1, \ldots, C_n\}$ is an orthonormal basis for $\mathbb{C}^n(\mathbb{R}^n)$

Definition.

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

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

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

 $\therefore Q$ is unitary matrix.

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

... There exists an orthonormal basis consisting of eigenvectors of T

Definition (Projection **T**). $\mathbf{W}_1, \mathbf{W}_2$ are subspaces of $\mathbf{V}, \mathbf{V} = \mathbf{W}_1 \bigoplus \mathbf{W}_2$, 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 $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \in \mathbf{W}_1 \bigoplus \mathbf{W}_2$.

Note. $im\mathbf{T} = \mathbf{W}_1 = \{\mathbf{x} \in \mathbf{V} \mid \mathbf{T}\mathbf{x} = \mathbf{x}\}$ and $\ker \mathbf{T} = \mathbf{W}_2$. $\therefore \mathbf{V} = im\mathbf{T} \bigoplus \ker \mathbf{T}$. 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 $V = W_1 \bigoplus W_2 = W_1 \bigoplus W_3$ does not imply that $W_2 = W_3$, W_1 does not uniquely determine T. However, for an orthogonal projection T, T is uniquely determined by its range.

Theorem. T is projection if and only if $T^2 = T$

Proof. $\Longrightarrow \mathbf{x} := \mathbf{x}_1 + \mathbf{x}_2 \in \mathbf{W}_1 \bigoplus \mathbf{W}_2$, 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 $\mathbf{x} - \mathbf{T}\mathbf{x} \in \ker \mathbf{T}$.

$$\mathbf{T}(\mathbf{x} - \mathbf{T}\mathbf{x}) = \mathbf{T}\mathbf{x} - \mathbf{T}^2\mathbf{x}$$

$$= \mathbf{T}\mathbf{x} - \mathbf{T}\mathbf{x}$$

$$= \mathbf{0}$$

$$\therefore \mathbf{x} - \mathbf{T}\mathbf{x} \in \ker \mathbf{T}$$

Now we want to show that $\operatorname{im} \mathbf{T} \cap \ker \mathbf{T} = \mathbf{O}$. Let $\mathbf{x} \in \operatorname{im} \mathbf{T} \cap \ker \mathbf{T}$ be given. Since $\mathbf{x} \in \ker \mathbf{T}$, $\mathbf{T}\mathbf{x} = \mathbf{0}$. Since $\mathbf{x} \in \operatorname{im} \mathbf{T}$, $\exists \mathbf{x}_0 \in \mathbf{V}$ 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 $\operatorname{im} \mathbf{T} \cap \ker \mathbf{T} = \mathbf{0}$, i.e., $\mathbf{V} = \operatorname{im} \mathbf{T} \bigoplus \ker \mathbf{T}$. Put $\mathbf{W}_1 := \operatorname{im} \mathbf{T}$, $\mathbf{W}_2 := \ker \mathbf{T}$. $\therefore \mathbf{T}$ is projection on \mathbf{W}_1

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

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

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

$$\langle \mathbf{x}, \mathbf{T} \mathbf{y} \rangle = \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{T} (\mathbf{y}_1 + \mathbf{y}_2)$$

$$= \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{T} \mathbf{y}_1 \rangle$$

$$(\because \mathbf{y}_2 \in \ker \mathbf{T})$$

$$= \langle \mathbf{x}_1, \mathbf{T} \mathbf{y}_1 \rangle + \langle \mathbf{x}_2, \mathbf{T} \mathbf{y}_1 \rangle$$

$$= \langle \mathbf{x}_1, \mathbf{y}_1 \rangle + \langle \mathbf{x}_2, \mathbf{y}_1 \rangle$$

$$(\because \mathbf{y}_1 \in \operatorname{im} \mathbf{T}, \mathbf{T}^2 = \mathbf{T}, \mathbf{y}_1 = \mathbf{T} \mathbf{y}' = \mathbf{T}^2 \mathbf{y}' = \mathbf{T} \mathbf{y}_1)$$

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

$$\langle \mathbf{T}\mathbf{x}, \, \mathbf{y} \rangle = \langle \mathbf{T}\mathbf{x}_1 + \mathbf{T}\mathbf{x}_2, \, \mathbf{y}_1 + \mathbf{y}_2 \rangle$$

$$= \langle \mathbf{T}\mathbf{x}_1, \, \mathbf{y}_1 + \mathbf{y}_2 \rangle$$

$$= \langle \mathbf{x}_1, \, \mathbf{y}_1 \rangle + \langle \mathbf{x}_1, \, \mathbf{y}_2 \rangle$$

$$= \langle \mathbf{x}_1, \, \mathbf{y}_1 \rangle$$

$$\langle \mathbf{x}, \, \mathbf{T}\mathbf{y} \rangle = \langle \mathbf{T}\mathbf{x}, \, \mathbf{y} \rangle$$

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

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

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

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle$$

$$= \langle \mathbf{T}^*\mathbf{x}, \mathbf{y} \rangle$$

$$= \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle$$

$$= \langle \mathbf{x}, \mathbf{0} \rangle$$

$$= 0$$

$$\therefore \mathbf{x} \in \ker \mathbf{T}^{\perp}$$

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

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

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

$$\langle \mathbf{T}\mathbf{y}, \, \mathbf{y} - \mathbf{T}\mathbf{y} \rangle = \langle \mathbf{y}, \, \mathbf{T}^*(\mathbf{y} - \mathbf{T}\mathbf{y}) \rangle$$

$$= \langle \mathbf{y}, \, \mathbf{T}(\mathbf{y} - \mathbf{T}\mathbf{y}) \rangle$$

$$= \langle \mathbf{y}, \, \mathbf{0} \rangle$$

$$= 0$$

$$\therefore \mathbf{y} = \mathbf{T}\mathbf{y}$$

Then we want to show that $\ker \mathbf{T} = \mathrm{im} \mathbf{T}^{\perp}$. $\ker \mathbf{T} \subset (\ker \mathbf{T}^{\perp})^{\perp} = \mathrm{im} \mathbf{T}^{\perp}$. Let $\mathbf{x} \in \mathrm{im} \mathbf{T}^{\perp}$. $\forall \mathbf{y} \in \mathbf{V}$,

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

Thus $\mathbf{x} \in \ker \mathbf{T}$, $\ker \mathbf{T} = \operatorname{im} \mathbf{T}^{\perp}$

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

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

Theorem (Spectral Theorem). **V** is finite-dimensional vector space over **F** and **T** is linear operator on **V** with distinct eigenvalue $\lambda_1, \ldots, \lambda_k$. **T** is normal if $\mathbf{F} = \mathbb{C}$ or **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). $\mathbf{V} = \mathbf{W}_1 \oplus \cdots \oplus \mathbf{W}_k$ (eigenspace decomposition)

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

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

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

(e).
$$\mathbf{T} = \lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k$$
 (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 **T** is normal or self-adjoint, two eigenvectors corresponding each to two different eigenvalues are orthogonal. $\mathbf{W}'_i \subset \mathbf{W}_i^{\perp}$

$$\dim \mathbf{W}_i' = \sum_{i \neq j} \dim \mathbf{W}_j$$

$$= \dim \mathbf{V} - \dim \mathbf{W}_i$$

$$= \dim \mathbf{W}_i^{\perp}$$

$$\therefore \mathbf{W}_i' = \mathbf{W}_i^{\perp}$$

(c). Proof. $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_k$ where $\mathbf{x}_i \in \mathbf{W}_i$. $\mathbf{T}_i \mathbf{T}_j(\mathbf{x}) = \mathbf{T}_i \mathbf{x}_j$ $\begin{cases} i = j & \mathbf{T}_i \mathbf{x}_j = \mathbf{x}_i = \mathbf{T}_i \mathbf{x} \\ i \neq j & \mathbf{T}_i \mathbf{x}_j = \mathbf{0} \\ \therefore \mathbf{T}_i \mathbf{T}_j = \delta_{ij} \mathbf{T}_i \end{cases}$

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

$$\mathbf{x} = \mathbf{x}_i + (\mathbf{x}_1 + \dots + \mathbf{x}_{i-1} + \mathbf{x}_{i+1} + \dots + \mathbf{x}_k)$$

$$= \mathbf{T}_i \mathbf{x} + (\mathbf{T}_1 \mathbf{x} + \dots + \mathbf{T}_{i-1} \mathbf{x} + \mathbf{T}_{i+1} \mathbf{x} + \dots + \mathbf{T}_k \mathbf{x})$$

$$= (\mathbf{T}_1 + \dots + \mathbf{T}_k) \mathbf{x}$$

$$\therefore \mathbf{I} = \mathbf{T}_1 + \dots + \mathbf{T}_k$$

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

$$\mathbf{Tx} = \mathbf{T}(\mathbf{x}_1 + \dots + \mathbf{x}_k)$$

$$= \mathbf{Tx}_1 + \dots + \mathbf{Tx}_k$$

$$= \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k$$

$$= \lambda_1 \mathbf{T}_1 \mathbf{x} + \dots + \lambda_k \mathbf{T}_k \mathbf{x}$$

$$= (\lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k) \mathbf{x}$$

$$\therefore = \lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k$$

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

Note. If T is normal(self-adjoint) and β is the union of orthonormal bases of the \mathbf{W}_i 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. T is unitary \iff T is normal and $|\lambda_i| = 1$ for i = 1, ..., k

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

$$\mathbf{T}\mathbf{T}^*\mathbf{v}_i = \mathbf{T}(\overline{\lambda}_i\mathbf{v}_i)$$

$$= |\lambda_i|^2\mathbf{v}_i$$

$$= \mathbf{v}_i \ (\because \mathbf{T}\mathbf{T}^* = \mathbf{I})$$

$$\therefore |\lambda_i|^2 = 1$$

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

$$\mathbf{T}^*\mathbf{T} = (\overline{\lambda}_1 \mathbf{T}_1 + \dots + \overline{\lambda}_k \mathbf{T}_k)(\lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k)$$

$$= |\lambda_1|^2 \mathbf{T}_1^2 + \dots + |\lambda_k|^2 \mathbf{T}_k^2$$

$$= \mathbf{T}_1 + \dots + \mathbf{T}_k$$

$$= \mathbf{I}$$

∴ **T** is unitary

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

Proof. \Longrightarrow Suppose that **T** is self-adjoint. For all i

$$\lambda_{i}\mathbf{x}_{i} = \mathbf{T}\mathbf{x}_{i}$$

$$= \mathbf{T}^{*}\mathbf{x}_{i}$$

$$= \overline{\lambda}_{i}\mathbf{x}_{i}$$

$$\therefore \overline{\lambda}_{i} = \lambda_{i}$$

$$\therefore \lambda_{i} \in \mathbb{R}$$

 \leftarrow

$$\mathbf{T} = \lambda_1 \mathbf{T}_1 + \lambda_2 \mathbf{T}_2 + \dots + \lambda_k \mathbf{T}_k$$

$$\mathbf{T}^* = \overline{\lambda}_1 \mathbf{T}_1 + \dots + \overline{\lambda}_k \mathbf{T}_k$$

$$= \lambda_1 \mathbf{T}_1 + \dots + \lambda_k \mathbf{T}_k$$

$$= \mathbf{T}$$

 \therefore T is self-adjoint

Definition (Singular Value Decomposition). **V**, **W** are finite-dimensional inner product space. **T**: $\mathbf{V} \longrightarrow \mathbf{W}$: 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 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ such that

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

 $\begin{pmatrix} 0 & (i > r) \\ Conversely & if the preceding conditions are satisfied, then \end{pmatrix}$

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

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