FALL 2021 Linear Algebra Discussion 5

Before we start, let us recap important definitions.

Definition .1. Given a point $\mathbf{y} \in \mathcal{V}$, and a subspace $\mathcal{S} \subset \mathcal{V}$, a vector $\widehat{\mathbf{y}}_S$ is called an orthogonal projection of $\mathbf{y}$ onto $\mathcal{S}$ if

- $\widehat{\mathbf{y}}_S \in \mathcal{S}$
- $\langle \mathbf{y} - \widehat{\mathbf{y}}_S, \mathbf{v} \rangle = 0 \quad \forall \mathbf{v} \in \mathcal{S}$

Let us define the following quantities. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for $\mathcal{S}$. Then

$$\mathbf{A} = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \ldots & \langle \mathbf{v}_k, \mathbf{v}_1 \rangle \\ \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \ldots & \langle \mathbf{v}_k, \mathbf{v}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{v}_1, \mathbf{v}_k \rangle & \langle \mathbf{v}_2, \mathbf{v}_k \rangle & \ldots & \langle \mathbf{v}_k, \mathbf{v}_k \rangle \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} \langle \mathbf{y}, \mathbf{v}_1 \rangle \\ \langle \mathbf{y}, \mathbf{v}_2 \rangle \\ \vdots \\ \langle \mathbf{y}, \mathbf{v}_k \rangle \end{bmatrix} \tag{1}$$

We showed in class that the orthogonal projection exists if and only if the Normal Equation $\mathbf{b} = \mathbf{A}\mathbf{x}$ has a solution. Moreover, the orthogonal projection is given by

$$\widehat{\mathbf{y}}_S = \sum_{i=1}^{k} x_i \mathbf{v}_i$$

Question 1: The Normal Equation $\mathbf{b} = \mathbf{A}\mathbf{x}$ for given $\mathbf{A} \in \mathbb{F}^{k\times k}$ and $\mathbf{b} \in \mathbb{F}^{k\times 1}$ ($\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$) in Eq(1) has always a solution. Hence, the orthogonal projection always exists.

Solution 1: We will show that $\text{rank}(\mathbf{A}) = k$. To do so, we will argue $N(\mathbf{A}) = \{\mathbf{0}\}$. Suppose $\mathbf{z} \in N(\mathbf{A})$. Then $\mathbf{A}\mathbf{z} = \mathbf{0}$, which in turn implies $\mathbf{z}^H\mathbf{A}\mathbf{z} = \mathbf{0}$. Note that $A_{i,j} = \langle \mathbf{v}_j, \mathbf{v}_i \rangle$. Using the properties of the inner product, we can write this product as

$$0 = \mathbf{z}^H\mathbf{A}\mathbf{z} = \sum_{i=1}^{k}\sum_{j=1}^{k} z_i^* A_{ij} z_j = \sum_{i=1}^{k}\sum_{j=1}^{k} z_i^* z_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \sum_{i=1}^{k}\sum_{j=1}^{k} \langle z_j\mathbf{v}_j, z_i\mathbf{v}_i \rangle$$

$$= \sum_{i=1}^{k} \langle \sum_{j=1}^{k} z_j\mathbf{v}_j, z_i\mathbf{v}_i \rangle = \langle \sum_{j=1}^{k} z_j\mathbf{v}_j, \sum_{i=1}^{k} z_i\mathbf{v}_i \rangle = \langle \sum_{i=1}^{k} z_j\mathbf{v}_i, \sum_{i=1}^{k} z_i\mathbf{v}_i \rangle$$

This implies $\sum_{i=1}^{k} z_i\mathbf{v}_i = \mathbf{0}$. Since vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent, $\mathbf{z} = \mathbf{0}$. Hence we showed $N(\mathbf{A}) = \{\mathbf{0}\}$. This in turn shows $\text{rank}(\mathbf{A}) = k$, and hence there always exists $\mathbf{x} \in \mathbb{F}^k$ for which $\mathbf{b} = \mathbf{A}\mathbf{x}$.

Question 2: Show that if $\mathbf{Q} \in \mathbb{C}^{m\times n}$, then $\text{rank}(\mathbf{Q^H Q}) = \text{rank}(\mathbf{Q})$.

Solution 2: We will show $N(\mathbf{Q}) = N(\mathbf{Q^H Q})$. We will start with $\mathbf{x} \in \mathbf{Q}$.

$$\mathbf{x} \in N(\mathbf{Q}) \implies \mathbf{Q}\mathbf{x} = \mathbf{0} \implies \mathbf{Q^H Q x} = \mathbf{0} \implies \mathbf{x} \in N(\mathbf{Q^H Q})$$

Now we will consider $\mathbf{x} \in \mathbf{Q^H Q}$.

$$\mathbf{x} \in N(\mathbf{Q^H Q}) \implies \mathbf{Q^H Q x} = \mathbf{0} \implies \mathbf{x^H Q^H Q x} = 0 \implies \|\mathbf{Q}\mathbf{x}\|_2^2 = 0 \implies \mathbf{Q}\mathbf{x} = \mathbf{0} \implies \mathbf{x} \in N(\mathbf{Q})$$

Hence $N(\mathbf{Q}) = N(\mathbf{Q^H Q})$. Note that $\mathbf{Q^H Q} \in \mathbb{C}^{m\times m}$. We will use the rank-nullity theorem to see

$$\text{rank}(\mathbf{Q}) = n - \dim(N(\mathbf{Q})) = n - \dim(N(\mathbf{Q^H Q})) = \text{rank}(\mathbf{Q^H Q})$$

This completes the proof.

Question 3: Suppose $\mathbf{y} \in \mathbb{C}^m$ and $\mathcal{S} \subset \mathbb{C}^m$ be a subspace. We define the standard inner product in $\mathbb{C}^m$ as $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_2^H \mathbf{v}_1$. Let $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k] \in \mathbb{C}^{m \times k}$ be the matrix of basis vector of $\mathcal{S}$. Then show that the normal equation is equivalent to

$$\mathbf{V^H y} = \mathbf{V^H V x}$$

Find the projection of $\mathbf{y}$ onto $\mathcal{S}$.

Solution 3: We will show $\mathbf{A}$ and $\mathbf{b}$ defined in (1) is equal to $\mathbf{V}^H\mathbf{V}$ and $\mathbf{V^H y}$. First see that $(\mathbf{A})_{i,j} = \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \mathbf{v}_i^H \mathbf{v}_j$. Then $(\mathbf{V}^H\mathbf{V})_{i,j} = \mathbf{v}_j^H \mathbf{v}_i$. Hence this shows

$$\mathbf{A} = \mathbf{V}^H \mathbf{V}.$$

Similarly

$$(\mathbf{b})_i = \langle \mathbf{y}, \mathbf{v}_i \rangle = \mathbf{v}_i^H \mathbf{y} = (\mathbf{V^H y})_i.$$

Hence $\mathbf{b} = \mathbf{V^H y}$. This shows the normal equation is equivalent to

$$\mathbf{V^H y} = \mathbf{V^H V x}.$$

Since $\mathbf{V}$ consists of basis vectors, $\text{rank}(\mathbf{V}) = k$. Then $\text{rank}(\mathbf{V}^H\mathbf{V}) = k$. This show $\mathbf{V}^H\mathbf{V}$ is invertible. Hence the solution to the Normal Equation is given as

$$\mathbf{x} = (\mathbf{V}^H\mathbf{V})^{-1}\mathbf{V}^H\mathbf{y}.$$

The projection of $\mathbf{y}$ onto $\mathcal{S}$ is given as

$$\widehat{\mathbf{y}}_S = \sum_{i=1}^{k} x_i \mathbf{v}_i = \mathbf{V}\mathbf{x} = \mathbf{V}(\mathbf{V}^H\mathbf{V})^{-1}\mathbf{V}^H\mathbf{y}$$

Question 4: Let $\mathcal{S} = \{\mathbf{X} \in \mathbb{R}^{m \times m},\ X_{i,j} = X_{j,i}\}$ be the set of symmetric matrices. We define the inner product in $\mathbb{R}^{m \times m}$ as $\langle \mathbf{S}_1, \mathbf{S}_2 \rangle = \text{trace}(\mathbf{S}_2^T \mathbf{S}_1)$.

- Show that the orthogonal projection of any matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ onto $\mathcal{S}$ is

$$\frac{\mathbf{A} + \mathbf{A}^T}{2}$$

Solution 4: Note that in Discussion 2, we showed that $\mathcal{S}$ is a subspace of $\mathbb{R}^{m \times m}$. Hence the orthogonal projection is well defined. Now we can easily show that

$$(\frac{\mathbf{A} + \mathbf{A}^T}{2})^T = \frac{\mathbf{A} + \mathbf{A}^T}{2} \implies \frac{\mathbf{A} + \mathbf{A}^T}{2} \in \mathcal{S}$$

In discussion 2, we constructed the basis of $\mathcal{S}$. Remember we defined the matrix $\mathbf{E}_{i,j}$ for given $1 \leq i, j \leq m$:

$$(\mathbf{E}_{i,j})_{k,l} = \begin{cases} 1 & \text{if } i = k, j = l \text{ or } i = l, j = k \\ 0 & \text{otherwise} \end{cases}$$

Based on this definition, we showed the set

$$\mathcal{A} = \{\mathbf{E}_{i,j} \mid 1 \leq j \leq i \leq m\}$$

is a basis of $\mathcal{S}$. We need to show

$$\langle \mathbf{A} - \frac{\mathbf{A} + \mathbf{A}^T}{2}, \mathbf{S} \rangle = 0, \quad \forall \mathbf{S} \in \mathcal{S}$$

This condition is equivalent to

$$\langle \mathbf{A} - \frac{\mathbf{A}+\mathbf{A}^T}{2}, \mathbf{E}_{i,j} \rangle = 0, \quad \forall \mathbf{E}_{i,j} \in \mathcal{A}$$

We evaluate the inner product as follows:

$$\begin{aligned} \langle \mathbf{A} - \frac{\mathbf{A}+\mathbf{A}^T}{2}, \mathbf{E}_{i,j} \rangle &= \langle \frac{\mathbf{A}-\mathbf{A}^T}{2}, \mathbf{E}_{i,j} \rangle = \text{trace}(\mathbf{E}_{i,j}^T(\frac{\mathbf{A}-\mathbf{A}^T}{2})) \\ &= \sum_{k=1}^{m}\sum_{l=1}^{m}(\mathbf{E}_{i,j})_{k,l}(\frac{\mathbf{A}-\mathbf{A}^T}{2})_{k,l} = (\frac{\mathbf{A}-\mathbf{A}^T}{2})_{i,j} + (\frac{\mathbf{A}-\mathbf{A}^T}{2})_{j,i} \\ &= \frac{\mathbf{A}_{i,j} - \mathbf{A}_{j,i}^T + \mathbf{A}_{j,i} - \mathbf{A}_{i,j}^T}{2} = \frac{\mathbf{A}_{i,j} - \mathbf{A}_{i,j} + \mathbf{A}_{j,i} - \mathbf{A}_{j,i}}{2} = 0 \end{aligned}$$

Hence $\frac{\mathbf{A}+\mathbf{A}^T}{2}$ is the orthogonal projection of $\mathbf{A}$ onto $\mathcal{S}$.