Unit 4(b) Lecture Notes for MAT224

Xuanqi Wei 1009353209

7-9 March 2023

Let V be a vector space. An *inner product* on V is any mapping $<\,,\,>: V \times V \to R$ with these three properties:

(i) For all vectors x, y, z, $\in V$ and all $c \in R$ we get that $< cx + y, z >= c < x, z > + < y, z >$.

(ii) For all x, $y \in V$, $< x, y >=< y, x >$

(iii) For all $x \in V$, $< x, x >\geq 0$, and $< x, x >= 0$ only if $x = 0$.

Show that the dot product of two vectors in $\mathbb{R}^2$ is an inner product. In other words, prove that the following is an inner product:

$$\left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\rangle = (u_1) \cdot (v_1) + (u_2) \cdot (v_2)$$

1. Linearity in the first argument:

$$\left\langle c\mathbf{u} + \mathbf{v}, \mathbf{w} \right\rangle = (cu_1 + v_1)w_1 + (cu_2 + v_2)w_2$$
$$= cu_1 w_1 + cu_2 w_2 + v_1 w_1 + v_2 w_2$$
$$= c\left\langle \mathbf{u}, \mathbf{w} \right\rangle + \left\langle \mathbf{v}, \mathbf{w} \right\rangle$$

2. Conjugate symmetry:

$$\overline{\left\langle \mathbf{v}, \mathbf{u} \right\rangle} = \overline{v_1 u_1 + v_2 u_2}$$
$$= \overline{v_1 u_1} + \overline{v_2 u_2}$$
$$= \overline{v_1} \cdot \overline{u_1} + \overline{v_2} \cdot \overline{u_2}$$
$$= u_1 \cdot v_1 + u_2 \cdot v_2$$
$$= \left\langle \mathbf{u}, \mathbf{v} \right\rangle$$

3. Finally, we need to show that the inner product satisfies the positive definite property, which means that $\langle v, v \rangle > 0$ for all $v \in \mathbb{R}^2$ except for $v = \mathbf{0}$, and $\langle \mathbf{0}, \mathbf{0} \rangle = 0$. This follows immediately from the fact that the dot product of any non-zero vector with itself is always positive, and the dot product of the zero vector with any other vector is always zero.

Give another example an inner product.
An example of another inner product is the standard inner product on $\mathbb{R}^n$ given by: $< x, y >= x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$ for vectors $\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \vdots & x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 & y_2 & \vdots & y_n \end{pmatrix}$.

§1 4.4 Orthogonal Projections and Gram-Schmidt Process

Define "orthogonal set" and "orthonormal set". What is the difference?

An **orthogonal set** of vectors in a vector space V is a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ such that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$, where $\cdot$ denotes the inner product on V.

An **orthonormal set** of vectors in a vector space V is a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ such that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$, and $|\mathbf{v}_i| = 1$ for all i, where $\cdot$ denotes the inner product on V and $|\cdot|$ denotes the norm induced by the inner product.

The difference between an orthogonal set and an orthonormal set is that the latter is a set of vectors with length 1 (i.e., normalized), while the former is not necessarily normalized.

Give an example of a finite dimensional vector space V, a subspace $W \subseteq V$, and its orthogonal compliment in V, called $W^\perp$.

Consider the vector space $V = \mathbb{R}^3$ with the standard inner product $\langle \cdot, \cdot \rangle$. Let W be the subspace spanned by the vectors $(1, 0, 1)$ and $(0, 1, 1)$, i.e. $W = \operatorname{span}(1, 0, 1), (0, 1, 1)$. Then, we can find a basis for $W^\perp$ by solving the system of equations

$$\langle (1, 0, 1), (x, y, z) \rangle = 0$$
$$\langle (0, 1, 1), (x, y, z) \rangle = 0$$

which yields the equation $x + z = 0$ and $y + z = 0$. Hence, a basis for $W^\perp$ is given by the vector $(1, -1, 0)$. Therefore, $W^\perp = \operatorname{span}(1, -1, 0)$.

Choose an element $v \in V$ from above. How many ways are there to write $v = w_1 + w_2$ where $w_1 \in W$ and $w_2 \in W^\perp$? (Hint: See Proposition 4.4.3(d))

Given a subspace W and its orthogonal complement $W^\perp$, any vector $v \in V$ can be uniquely expressed as $v = w_1 + w_2$ where $w_1 \in W$ and $w_2 \in W^\perp$.

To see why, consider the projection of v onto W given by $w_1 = \operatorname{proj}_W(v)$. Then, $w_2 = v - w_1$ is the projection of v onto $W^\perp$. Since w_1 is in W and w_2 is in $W^\perp$, we have $v = w_1 + w_2$.

To see that this expression is unique, suppose there exist $w_1', w_2' \in V$ such that $v = w_1 + w_2 = w_1' + w_2'$ with $w_1, w_1' \in W$ and $w_2, w_2' \in W^\perp$. Then, $w_1 - w_1' = w_2' - w_2 \in W \cap W^\perp$. Since $W \cap W^\perp = 0$, we must have $w_1 = w_1'$ and $w_2 = w_2'$. Thus, the expression for v as the sum of a vector in W and a vector in $W^\perp$ is unique.

Let α be a basis of subspace W and β be a basis of its orthogonal complement $W^\perp$ in vector space V. How can we create a basis of V from α and β?

To create a basis of V from α and β, we can simply concatenate the two bases. That is, if $\alpha = v_1, v_2, \ldots, v_k$ and $\beta = w_1, w_2, \ldots, w_m$, then $\gamma = v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_m$ is a basis of V.

To show that γ is a basis, we need to show that it spans V and is linearly independent.

First, since W and $W^\perp$ together span V, every vector in V can be written as a linear combination of the vectors in α and β. Therefore, γ spans V.

Next, we need to show that γ is linearly independent. Suppose there exist scalars $a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_m$ such that

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k + b_1 w_1 + b_2 w_2 + \ldots + b_m w_m = 0$$

Since α is a basis of W, we know that $a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = \mathbf{0}$ implies that $a_1 = a_2 = \cdots = a_k = 0$. Similarly, since β is a basis of $W^\perp$, we know that $b_1 w_1 + b_2 w_2 + \cdots + b_m w_m = \mathbf{0}$ implies that $b_1 = b_2 = \cdots = b_m = 0$. Therefore, the only way the linear combination $a_1 v_1 + a_2 v_2 + \cdots + a_k v_k + b_1 w_1 + b_2 w_2 + \cdots + b_m w_m = \mathbf{0}$ can hold is if $a_1 = a_2 = \cdots = a_k = b_1 = b_2 = \cdots = b_m = 0$. Therefore, γ is linearly independent, and hence is a basis of V.

Let $\vec{a}$, $\vec{b}$, and $\vec{c} \in \mathbb{R}^2$. If $\vec{a} + \vec{b} = \vec{c}$ then $\vec{c} - \vec{b} \in span\{\vec{a}\}$ What has to be true of $\vec{a}$ and $\vec{b}$ in order to conclude that $\vec{a}$ is the orthogonal projection of $\vec{c}$ onto $span\{\vec{a}\}$?

1. $\vec{a}$ must be a non-zero vector, i.e., $\vec{a} \neq \vec{0}$.

2. $\vec{c} - \vec{a}$ must be orthogonal to $span\,\vec{a}$, i.e., $\vec{c} - \vec{a}$ is perpendicular to $\vec{a}$.

Write a formula for $P_W(x)$, the projection of vector $x \in V$ onto the subspace W of V.

If $\alpha = w_1, w_2, \ldots, w_k$ is an orthogonal basis for W, then the projection of x onto W is given by

$$P_W(x) = \frac{<x, w_1>}{\|w_1\|^2} w_1 + \frac{<x, w_2>}{\|w_2\|^2} w_2 + \ldots + \frac{<x, w_k>}{\|w_k\|^2} w_k$$

If α is an orthonormal basis for W, then the formula is simply

$$P_W(x) = <x, w_1> w_1 + <x, w_2> w_2 + \ldots + <x, w_k> w_k$$

We know that for any subspace W of $\mathbb{R}^n$, there exists an orthonormal basis of W from Theorem 4.4.9. Give an example of a set of three vectors $\{u_1, u_2, u_3\}$ and show how to create an **orthogonal** set $\{v_1, v_2, v_3\}$ so that $v_1 = u_1$, using the Gram-Schmidt process.

What extra step is required to turn $\{v_1, v_2, v_3\}$ into an **orthonormal** set?

Let $u_1 = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$, $u_3 = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}$.

We begin by setting $v_1 = u_1$. Then, for v_2, we subtract off the projection of u_2 onto v_1:

$$v_2 = u_2 - proj_{v_1}(u_2)$$

where $proj_{v_1}(u_2) = \frac{\langle u_2, v_1 \rangle}{|v_1|^2} v_1$. Since $v_1 = u_1$ and $|u_1| = \sqrt{2}$, we have:

$$proj_{v_1}(u_2) = \frac{< u_2, v_1 >}{\|v_2\|^2} v_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Therefore,

$$v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Next, for v_3, we subtract off the projections of u_3 onto v_1 and v_2:

$$v_3 = u_2 - proj_{v_1}(u_3) - proj_{v_2}(u_3)$$

We have already computed $proj v_1(u_3)$, so we only need to compute $proj v_2(u_3)$. Since v_2 is already normalized, we have:

$$proj v_2(u_3) = < v_3, v_2 > v_2 = \frac{3}{2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Therefore,

$$v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

Thus, the orthogonal set v_1, v_2, v_3 is:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

§2   4.5 Symmetric Matrices

Let A be the $n \times n$ matrix $\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$.

Another way to write the same thing is to say that

$$A = [a_{ij}] \qquad (*)$$

for $i, j \in \{1, 2, 3, \ldots n\}$. Write a similar expression to $(*)$ for A^t, the transpose of A.

$$A = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{11} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

Describe in words what we know about the ij^{th} entry of A, if A is a **symmetric** matrix?

If A is a symmetric matrix, we know that its ij^{th} entry is equal to its ji^{th} entry, meaning that $A_{ij} = A_{ji}$. In other words, the matrix is symmetric about its main diagonal. This implies that the entries above the main diagonal are reflections of the entries below the main diagonal, and vice versa. This can be thought of as a mirror image reflection along the main diagonal.

Proposition 4.5.2(a) and Corollary 4.5.2(b) relate inner products to symmetric matrices. Give the definition of a "symmetric linear transformation" and explain how 4.5.2(a)-(b) allows us to state this definition.

A linear transformation $T : V \to V$ is called symmetric if $\langle T(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, T(\mathbf{w}) \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$, where $\langle \cdot, \cdot \rangle$ is the inner product on V. In other words, a linear transformation is symmetric if the inner product of the output of the transformation applied to one vector with another vector is equal to the inner product of that vector with the output of the transformation applied to the other vector.

Proposition 4.5.2(a) states that every inner product on a finite-dimensional real vector space V can be expressed as the standard inner product $\langle \cdot, \cdot \rangle_A$ induced by a unique positive definite symmetric matrix A. This means that we can identify the inner product on V with a symmetric matrix A.

Corollary 4.5.2(b) then tells us that for any symmetric matrix A, there exists a unique linear transformation $T : V \to V$ such that A is the matrix representation of T with respect to some orthonormal basis of V. This means that we can identify a symmetric matrix A with a symmetric linear transformation T on V. In other words, if A is a symmetric matrix, then there exists a symmetric linear transformation T such that A is the matrix representation of T with respect to some orthonormal basis of V.

What do we know about the eigenvalues of an $n \times n$ symmetric matrix?
We know that the eigenvalues of an $n \times n$ symmetric matrix A are all real. This is a consequence of the Spectral Theorem for symmetric matrices. Furthermore, we know that eigenvectors corresponding to distinct eigenvalues are orthogonal to each other.

If a polynomial $p(x)$ has n roots, counting multiplicities, what is the degree of $p(x)$?
If a polynomial $p(x)$ has n roots, counting multiplicities, then the degree of $p(x)$ is at least n. Specifically, the degree of $p(x)$ is equal to n if all n roots are distinct, and it is greater than n if there are repeated roots.

What does "counting multiplicities" mean?
Counting multiplicities means taking into account how many times each root appears as a solution of the polynomial equation. In other words, if a root appears more than once, we count it multiple times. For example, the polynomial $p(x) = (x - 2)^3(x + 1)$ has two distinct roots, $x = 2$ and $x = -1$, but 2 is a root with multiplicity 3, so we say that $p(x)$ has 4 roots counting multiplicities.

Let x_1 and x_2 be two vectors in $\mathbb{R}^n$ such that $x_1 \in E_{\lambda_1}$ and $x_2 \in E_{\lambda_2}$ where λ_1 and λ_2 are eigenvalues of an $n \times n$ matrix A. Prove that x_1 and x_2 are orthogonal.
Since $x_1 \in E_{\lambda_1}$, we know that $Ax_1 = \lambda_1 x_1$. Similarly, $Ax_2 = \lambda_2 x_2$ since $x_2 \in E_{\lambda_2}$. We can take the dot product of x_1 and x_2:

$$
\begin{aligned}
x_1 \cdot x_2 &= \left(\sum_{i=1}^{n} x_{1,i} x_{2,i} \right) \\
&= \lambda_1 \left(\sum_{i=1}^{n} x_{1,i} x_{2,i} \right) \\
&= \lambda_1 (x_1 \cdot x_2)
\end{aligned}
$$

Similarly, we can take the dot product of x_1 and Ax_2:

$$x_1 \cdot (A \cdot x_2) = \left(\sum_{i=1}^{n} x_{1,i}(A \cdot x_{2,i})\right)$$

$$= \lambda_2\left(\sum_{i=1}^{n} x_{1,i} x_{2,i}\right)$$

$$= \lambda_2(x_1 \cdot x_2)$$

Since A is a symmetric matrix, we know that $\lambda_1(x_1 \cdot x_2) = x_1 \cdot (Ax_2)$. Therefore, we have:

$$\lambda_1(x_1 \cdot x_2) = x_1 \cdot (Ax_2)$$

$$= x_2 \cdot (Ax_1)$$

$$= \lambda_1(x_1 \cdot x_2)$$

Since $\lambda_1 \neq \lambda_2$ (because x_1 and x_2 correspond to different eigenvalues), we must have $x_1 \cdot x_2 = 0$, i.e., x_1 and x_2 are orthogonal.