Chapter 10

Week10

10.1. Monday for MAT3040

10.1.1. Inner Product Space

- Symmetric: $F(\boldsymbol{u},\boldsymbol{w}) = F(\boldsymbol{w},\boldsymbol{u}), \forall \boldsymbol{u},\boldsymbol{w}$
- Non-degenerate: $F(\boldsymbol{u},\boldsymbol{w}) = 0, \forall \boldsymbol{w}$ implies $\boldsymbol{u} = \boldsymbol{0}$
- Positive definite: $F(\boldsymbol{v},\boldsymbol{v}) > 0, \forall \boldsymbol{v} \neq \boldsymbol{0}$

Classification. When we say V be a vector space over $\mathbb{F}$, we treat $\alpha \in \mathbb{F}$ as a scalar.

Definition 10.1 [Sesqui-linear Form] Let V be a vector space over $\mathbb{C}$. A **sesquilinear form** on V is a function $F : V \times V \to \mathbb{C}$ such that

1. $F(\boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}) = F(\boldsymbol{u},\boldsymbol{w}) + F(\boldsymbol{v},\boldsymbol{w})$
2. $F(\boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}) = F(\boldsymbol{u},\boldsymbol{v}) + F(\boldsymbol{u},\boldsymbol{w})$
3. $F(\overline{\lambda}\boldsymbol{v},\boldsymbol{w}) = F(\boldsymbol{v},\lambda\boldsymbol{w}) = \lambda F(\boldsymbol{v},\boldsymbol{w}), \forall \lambda \in \mathbb{C}$

In this case, we say F is **conjugate symmetric** if

$$F(\boldsymbol{v},\boldsymbol{w}) = \overline{F(\boldsymbol{w},\boldsymbol{v})}, \quad \forall \boldsymbol{v},\boldsymbol{w} \in V.$$

The definition for non-degenerateness, and positve definiteness is the same as that in bilinear form.

R In the sesquilinear form, why there is a $\bar{\lambda}$ shown in condition (3)?

Partial Answer: We want our F to be positive definite in many cases:

- Suppose that $F(\mathbf{v},\mathbf{v}) > 0$ and we do not have $\bar{\lambda}$ in sesquilinear form F, it follows that
$$F(i\mathbf{v},i\mathbf{v}) = i^2F(\mathbf{v},\mathbf{v}) = -F(\mathbf{v},\mathbf{v}) < 0$$
As a result, there will be no positive bilinear form for vector space over $\mathbb{C}$.

Therefore, $\bar{\lambda}$ is essential to guarantee that we have a positive definite form on vector space over $\mathbb{C}$, i.e.,

$$F(i\mathbf{v},i\mathbf{v}) = \bar{i}iF(\mathbf{v},\mathbf{v}) = F(\mathbf{v},\mathbf{v})$$

■ **Example 10.1** Consider $V = \mathbb{C}^n$, and a basic sesquilinear form is the Hermitian inner product:

$$F(\mathbf{v},\mathbf{u}) = \mathbf{v}^{\mathrm{H}}\mathbf{u} = \begin{pmatrix} \bar{v_1} & \cdots & \bar{v_n} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \sum_{i=1}^{n} \bar{v}_i w_i$$

In this case, we do not have symmetric property $F(\mathbf{v},\mathbf{w}) = F(\mathbf{w},\mathbf{v})$ any more, instead, we have the conjugate symmetric property $F(\mathbf{v},\mathbf{w}) = \overline{F(\mathbf{w},\mathbf{v})}$. ■

Definition 10.2 [Inner Product] A real (complex) vector space V with a bilinear (sesquilinear) form with symmetric (conjugate symmetric) and positive definite property is called an **inner product** on V. Any vector space equipped with inner product is called an **inner product space**. ■

Notation. We write $\langle\cdot,\cdot\rangle$ instead of $F(\cdot,\cdot)$ to denote inner product.

Definition 10.3 [Norm] The **norm** of a vector $\mathbf{v}$ is $\|\mathbf{v}\| = \sqrt{\langle\mathbf{v},\mathbf{v}\rangle}$. ■

R As a result, $\|\alpha\mathbf{v}\| = \sqrt{\langle\alpha\mathbf{v},\alpha\mathbf{v}\rangle} = \sqrt{\bar{\alpha}\alpha\langle\mathbf{v},\mathbf{v}\rangle} = \sqrt{|\alpha|^2\langle\mathbf{v},\mathbf{v}\rangle} = |\alpha|\|\mathbf{v}\|$.

The norm is well-defined since $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ (positive definiteness of inner product).

Definition 10.4 [Orthogonal] We say a family of vectors $S = \{\boldsymbol{v}_i \mid i \in I\}$ is **orthogonal** if

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0, \ \forall i \neq j$$

If furthermore $\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle = 1, \forall i$, then we say S is an **orthonormal** set. ■

R

1. The Cauchy-Scharwz inequality holds for inner product space:

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \|\boldsymbol{v}\|, \ \forall \boldsymbol{u}, \boldsymbol{v} \in V.$$

Proof. The proof for $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \in \mathbb{R}$ is the same as in MAT2040 course. Check Theorem (6.1) in the note

`https://walterbabyrudin.github.io/information/Notes/MAT2040.pdf`

However, for $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \in \mathbb{C} \setminus \mathbb{R}$, we need the re-scaling technique:

Let $\boldsymbol{w} = \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u}$, then $\langle \boldsymbol{w}, \boldsymbol{v} \rangle \in \mathbb{R}$:

$$\langle \boldsymbol{w}, \boldsymbol{v} \rangle = \langle \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u}, \boldsymbol{v} \rangle = \overline{\left(\frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \right)} \langle \boldsymbol{u}, \boldsymbol{v} \rangle = \frac{1}{\langle \boldsymbol{u}, \boldsymbol{v} \rangle} \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 1.$$

Applying the Cauchy-Scharwz inequality for $\langle \boldsymbol{w}, \boldsymbol{v} \rangle \in \mathbb{R}$ gives

$$\begin{aligned} \left| \langle \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u}, \boldsymbol{v} \rangle \right| &= |\langle \boldsymbol{w}, \boldsymbol{v} \rangle| \\ &\leq \|\boldsymbol{w}\| \|\boldsymbol{v}\| = \left\| \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u} \right\| \|\boldsymbol{v}\| \end{aligned}$$

Or equivalently,

$$\left| \frac{1}{\langle \boldsymbol{u}, \boldsymbol{v} \rangle} \right| |\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \left| \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \right| \|\boldsymbol{u}\| \|\boldsymbol{v}\|$$

Since $\left|\frac{1}{\langle \boldsymbol{u},\boldsymbol{v}\rangle}\right| = \left|\frac{1}{\overline{\langle \boldsymbol{u},\boldsymbol{v}\rangle}}\right|$, we imply

$$|\langle \boldsymbol{u},\boldsymbol{v}\rangle| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

■

2. The triangle inequality also holds for inner product process:

$$\|\boldsymbol{u}+\boldsymbol{v}\| \le \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$$

3. The Gram-Schmidt process holds for finite set of vectors: let $S = \{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ be (finite) linearly independent. Then we can construct an orthonormal set from S:

$$\boldsymbol{w}_1 = \boldsymbol{v}_1, \quad \boldsymbol{w}_{i+1} = \boldsymbol{v}_{i+1} - \frac{\langle \boldsymbol{v}_{i+1},\boldsymbol{w}_1\rangle}{\|\boldsymbol{w}_1\|^2} - \frac{\langle \boldsymbol{v}_{i+1},\boldsymbol{w}_2\rangle}{\|\boldsymbol{w}_2\|^2} - \cdots - \frac{\langle \boldsymbol{v}_{i+1},\boldsymbol{w}_i\rangle}{\|\boldsymbol{w}_i\|^2}, \; i = 1,\dots,n-1$$

Then after normalization, we obtain the constructed orthonormal set. Consequently, every finite dimensional inner product space has an orthonormal basis.

10.1.2. Dual spaces

Theorem 10.1 — Riesz Representation. Consider the mapping

$$\begin{aligned} \phi : \quad & V \to V^* \\ \text{with} \quad & \boldsymbol{v} \mapsto \phi_{\boldsymbol{v}} \\ \text{where} \quad & \phi_{\boldsymbol{v}}(w) = \langle \boldsymbol{v}, w\rangle, \; \forall w \in V \end{aligned}$$

Then the mapping ϕ is well-defined and it is an $\mathbb{R}$-linear transformation. Moreover, if V is finite dimensional, then ϕ is an isomorphism.

The $\mathbb{R}$-linear transformation $V \to V^*$ means that, when V, V^* are vector space over $\mathbb{R}$, the $\mathbb{R}$-linear transformation deduces into exactly the linear transformation.

R The $\mathbb{R}$-linear transformation $V \to V^*$ is **not** necessarily linear if V, V^* are vector spaces over $\mathbb{C}$.

However, we can transform a vector space over $\mathbb{C}$ into a vector space over $\mathbb{R}$:

- For example, suppose that $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis of V over $\mathbb{C}$, i.e.,

$$\boldsymbol{v} = \sum_{j=1}^{n} \alpha_j \boldsymbol{v}_j$$

where $\alpha_j = p_j + iq_j, \forall p_j, q_j \in \mathbb{R}$, then

$$\boldsymbol{v} = \sum_j p_j \boldsymbol{v}_j + \sum_j q_j(i\boldsymbol{v}_j), \; p_j, q_j \in \mathbb{R}$$

Therefore, $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, i\boldsymbol{v}_1, \ldots, i\boldsymbol{v}_n\}$ forms a basis of V over $\mathbb{R}$.

Note that $i\boldsymbol{v}_1$ cannot be considered as a linear combination of $\boldsymbol{v}_1$ over $\mathbb{R}$, but a linear combination of $\boldsymbol{v}_1$ over $\mathbb{C}$.

In particular, if $\phi : V \to V^*$ is a $\mathbb{R}$-linear transformation, then

$$\phi(i\boldsymbol{v}) \neq i\phi(\boldsymbol{v}), \text{ but } \phi(2\boldsymbol{v}) = 2\phi(\boldsymbol{v}).$$

Proof. 1. Well-definedness: We need to show $\phi_{\boldsymbol{v}} \in V^*$, i.e., for scalars a, b,

$$\phi_{\boldsymbol{v}}(a\boldsymbol{w}_1 + b\boldsymbol{w}_2) = \langle \boldsymbol{v}, a\boldsymbol{w}_1 + b\boldsymbol{w}_2 \rangle = a\langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle + b\langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle = a\phi_{\boldsymbol{v}}(\boldsymbol{w}_1) + b\phi_{\boldsymbol{v}}(\boldsymbol{w}_2)$$

Therefore, $\phi_{\boldsymbol{v}} \in V^*$.

2. $\mathbb{R}$-linearity of ϕ: it suffices to show

$$\phi(c\boldsymbol{v}_1 + d\boldsymbol{v}_2) = c\phi(\boldsymbol{v}_1) + d\phi(\boldsymbol{v}_2), \quad \forall c, d \in \mathbb{R}, \boldsymbol{v}_1, \boldsymbol{v}_2 \in V.$$

For all $\boldsymbol{w} \in V$, we have

$$\phi_{c\boldsymbol{v}_1 + d\boldsymbol{v}_2}(\boldsymbol{w}) = \langle c\boldsymbol{v}_1 + d\boldsymbol{v}_2, \boldsymbol{w} \rangle = c\langle \boldsymbol{v}_1, \boldsymbol{w} \rangle + d\langle \boldsymbol{v}_2, \boldsymbol{w} \rangle = c\phi_{\boldsymbol{v}_1}(\boldsymbol{w}) + d\phi_{\boldsymbol{v}_2}(\boldsymbol{w})$$

where the second equality holds because $c, d \in \mathbb{R}$.

Therefore,

$$\phi(c\boldsymbol{v}_1 + d\boldsymbol{v}_2) = c\phi(\boldsymbol{v}_1) + d\phi(\boldsymbol{v}_2).$$

■