Linear Algebra II

Felix Chen

Contents

We can introduce "angles" as well:

Definition 0.0.1 (Angles)**.** When $F = \mathbb{R}$, for $\alpha, \beta \in V \backslash \{0\}$, define

$$\angle(\alpha, \beta) = \arccos \frac{\langle \alpha, \beta \rangle}{\|\alpha\| \|\beta\|} \in [0, \pi].$$

We can see that $\alpha \perp \beta \iff \angle(\alpha, \beta) = \frac{\pi}{2}$.

When $F = \mathbb{C}$, the angle above can be complex, which doesn't make sense, so we won't talk about the angle in $\mathbb{C}$.

Definition 0.0.2 (Orthonormal basis)**.** Let V be an inner product space, let $S \subset V$ be a subset,

- If the vectors in S are pairwise orthogonal, we say S is an **orthogonal set**. Futhermore, if $\|\alpha\| = 1$ for all $\alpha \in S$, we say S is **orthonormal**.
- If S is a basis as well, then S is called an **orthogonal basis** or **orthonormal basis**, respectively.

Note that an orthogonal set can contain the zero vector.

Proposition 0.0.3

If S is an orthogonal set, and $0 \notin S$, then S is linearly independent.

Proof. Let $S = \{\alpha_1, \ldots, \alpha_n\}$, if

$$\sum_{j=1}^{n} c_j \alpha_j = 0,$$

take the inner product with α_j for $j = 1, \ldots, n$ we get $c_j = 0$, $\forall j$. □

Proposition 0.0.4

If $S = \{\alpha_1, \ldots, \alpha_m\}$ is an orthogonal set, then:

$$\left\| \sum_{j=1}^{m} \alpha_j \right\|^2 = \sum_{j=1}^{m} \|\alpha\|^2, \quad \left\langle \sum_{j=1}^{m} x_j \alpha_j, \sum_{j=1}^{m} y_j \alpha_j \right\rangle = \sum_{j=1}^{m} x_j \overline{y_j} \|\alpha_j\|^2.$$

Now we will prove the existence of orthogonal basis, We'll start from a basis $\{\beta_1, \beta_n\}$ to construct an orthogonal basis, and this process is called *Schmidt orthogonalization*.

Theorem 0.0.5 (Schmidt orthogonalization)
Let V be an n-dimensional inner product space, $\{\beta_1, \dots, \beta_n\}$ is a basis of V. Then there exists a unique orthogonal basis $\{\alpha_1, \dots, \alpha_n\}$, such that

$$(\beta_1, \dots, \beta_n) = (\alpha_1, \dots, \alpha_n)N,$$

where N is an upper triangular matrix with diagonal entries equal to 1.

Proof. The idea is to "project" β_j to the subspace spanned by $\beta_1, \dots, \beta_{j-1}$, and let α_j be the orthogonal part.

By induction, let $\beta_1 = \alpha_1$.

$$\alpha_j = \beta_j - \sum_{k=1}^{j-1} \frac{\langle \beta_j, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$

It's obvious that $\alpha_j \perp \alpha_k, \forall k = 1, \dots, j-1$, and $\operatorname{span}\{\alpha_1, \dots, \alpha_j\} = \operatorname{span}\{\beta_1, \dots, \beta_j\}$.

Thus $\{\alpha_1, \dots, \alpha_n\}$ is the desired orthogonal basis.

As for the uniqueness, actually α_j can be solved from β_j's: clearly $\alpha_1 = \beta_1$, and

$$\langle \beta_j, \alpha_k \rangle = N_{jk} \langle \alpha_k, \alpha_k \rangle \implies N_{jk} = \frac{\langle \beta_j, \alpha_k \rangle}{\|\alpha_k\|^2} \implies \alpha_j = \beta_j - \sum_{k=1}^{j-1} \frac{\langle \beta_j, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$

So α_j is uniquely determined by β_j's. □

Remark 0.0.6 — The above orthogonal basis can be converted to an orthonormal basis $\{\alpha_1', \dots, \alpha_n'\}$ s.t. N' is an upper triangular matrix with positive diagonal entries.

Corollary 0.0.7
Let $S \subset V \backslash \{0\}$ be orthogonal(-normal), then S can be extended to an orthogonal(-normal) basis.

Proposition 0.0.8
Let $S = \{\alpha_1, \dots, \alpha_m\} \subset V \backslash \{0\}$ be an orthogonal set, then for all $\beta \in \operatorname{span} S$ we have:

$$\beta = \sum_{k=1}^{m} \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$

Proposition 0.0.9 (Bessel's inequality)
Conditions as above, then $\forall \beta \in V$,

$$\sum_{k=1}^{m} \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2} \leq \|\beta\|^2.$$

Equality iff $\beta \in \operatorname{span} S$.

Proof. Complete S to an orthogonal basis, by previous propositions, the rest is trivial. □

Let $S \subset V$, define $S^\perp := \{\alpha \in V \mid \alpha \perp \beta, \forall \beta \in S\}$, $S^\perp$ is a vector space and $S^\perp = \operatorname{span}(S)^\perp$.

Proposition 0.0.10
Let V be a finite dimensional inner product space, $W \subset V$ is a subspace, we have $\dim W + \dim W^\perp = \dim V$.

Proof. Take an orthogonal basis B_1 of W, and complete it to an orthogonal basis B of V, then we claim that $B_2 := B \backslash B_1$ is a basis of $W^\perp$. Hence the conclusion follows. □

This means we always have $W \oplus W^\perp = V$.

The orthogonal completion is similar to the annihiltor we studied last semester, in fact, when we view $\langle \cdot, \beta \rangle$ as a function $f_\beta \in V^*$, $f_\beta \in S^0 \iff \beta \in S^\perp$. (Note that the inner product is linear with respect to only the first entry)

This process induces a map $\phi : V \to V^*$ by $\beta \mapsto f_\beta$. It's clear that ϕ is conjugate-linear. So ϕ is a linear map between *real* vector space $V \to V^*$, i.e. $\phi \in \operatorname{Hom}_{\mathbb{R}}(V, V^*)$. thus $\ker \phi = \{0\}$ implies ϕ is an isomorphism on $\mathbb{R}$, so ϕ is a bijection, $\phi(S^\perp) = S^0$.

For $E \subset V^*$, then $E^0 \subset V$, this corresponds to $\phi(S)^0 = S^\perp$. Indeed, $\alpha \in \phi(S)^0 \iff \forall \beta \in S, \langle \alpha, \beta \rangle = 0 \iff \alpha \in S^\perp$. Hence

$$\dim_{\mathbb{C}} W^\perp = 2 \dim_{\mathbb{R}} \phi(W^\perp) = 2 \dim_{\mathbb{R}} W^0 = \dim_{\mathbb{C}} W^0.$$

The above proposition can be derived directly by $\dim W + \dim W^0 = \dim V$.

We can also get $W = (W^0)^0 = \phi(W^\perp)^0 = (W^\perp)^\perp$.

Definition 0.0.11 (Orthogonal projection)**.** Since $V = W \oplus W^\perp$, for all $\alpha \in V$, there exists unique $\beta \in W, \gamma \in W^\perp$ s.t. $\alpha = \beta + \gamma$. Let $p_W : V \to W$ be the map $\alpha \mapsto \beta$, this is called the **orthogonal projection** from V to W.

§0.1 Adjoint maps

Let $\{\alpha_1, \ldots, \alpha_m\}$ be an orthonormal basis of W, then $p_W(\beta) = \sum_{j=1}^m \langle \beta, \alpha_j \rangle \alpha_j$. So p_W is a linear map. Moreover $p_W + p_{W^\perp} = \mathrm{id}_V$, $p_W^2 = p_W$. By our geometry intuition, $p_W \beta = \arg\min_\alpha \|\alpha - \beta\|$, this fact is useful in funtional analysis.

Recall that for $T \in L(V)$, $T^t \in L(V^*)$, then what's the map $\phi^{-1} \circ T^t \circ \phi$? Unluckily it's not T, but another map denoted by T^*, the **adjoint map** of T. Keep in mind that T^* depends on the inner product.

$$\begin{array}{ccc} V^* & \xrightarrow{T^t} & V^* \\ {\scriptstyle \phi}\uparrow & & {\scriptstyle \phi}\uparrow \\ V & \xrightarrow{T^*} & V \end{array}$$

Since $T^t \circ \phi = \phi \circ T^* \iff \langle T\alpha, \beta \rangle = \langle \alpha, T^*\beta \rangle$, $\forall \alpha, \beta \in V$, so T^* can be described as the map satisfying this relation.

Proposition 0.1.1

When $\mathcal{B}$ is an orthonormal basis, we have $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*$.

Proof. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, then $\phi(\mathcal{B})$ is the dual basis of $\mathcal{B}$. i.e. $\phi(\alpha_j)(\alpha_k) = \delta_{jk}$.

Hence $[T^t]_{\phi(\mathcal{B})} = [T]_{\mathcal{B}}^t$. Let $[T^*]_{\mathcal{B}} = A$, then

$$T^*\alpha_k = \sum_{j=1}^{n} A_{jk}\alpha_j \implies \phi(T^*\alpha_k) = \sum_{j=1}^{n} \overline{A_{jk}}\phi(\alpha_j).$$

So $[T^t]_{\phi(\mathcal{B})} = \overline{A}$, which completes the proof. □

Proposition 0.1.2

$\ker(T^*) = \mathrm{Im}(T)^{\perp}$, $\mathrm{Im}(T^*) = \ker(T)^{\perp}$. $(cT+U)^* = \bar{c}T^* + U^*$, $(TU)^* = U^*T^*$, $T^{**} = T$.

This means the map $T \mapsto T^*$ is a conjugate anti-automorphism of $L(V)$, and it's an involution.

If $T^* = T$, then we say T is **self-adjoint**, and if $T^* = -T$, we say T is **anti self-adjoint**.

Let $F = \mathbb{C}$, T is self-adjoint iff iT is anti self-adjoint. Like a function can be written as a sum of odd and even functions, $\forall T \in L(V)$, there exists unique self-adjoint T_1, T_2 s.t. $T = T_1 + iT_2$. In fact, $T_1 = \frac{T+T^*}{2}, T_2 = \frac{T-T^*}{2i}$.

Let $\mathcal{B}$ be an orthonormal basis, obviously T self-adjoint $\iff$ $[T]_{\mathcal{B}}$ Hermite.

Example 0.1.3

Let $W \subset V$, p_W be the orthogonal projection. then p_W is self-adjoint as we can choose an orthonormal basis $\mathcal{B}$, such that $[p_W]_{\mathcal{B}} = \mathrm{diag}\{I_k, 0\}$, where $k = \dim W$.

Let V, W be inner product spaces, we'll study the linear maps $T : V \to W$ which preserves the inner products, i.e.

$$\langle \alpha, \beta \rangle_V = \langle T\alpha, T\beta \rangle_W .$$

If T is an isomorphism, then we say T is the isomorphism between inner product spaces.

Proposition 0.1.4

T preserves inner product $\iff$ T is an isometry, i.e. preserves length.

In particular, isometry is always injective implies that inner product presering maps are always injective.

Proof. Trivial by polarization identity. □

Proposition 0.1.5

Let V, W be finite dimensional inner product spaces, $\dim V = \dim W$, $T \in \mathrm{Hom}(V, W)$, the followings are equivalent:

(1) T preserves inner product;

(2) T is an isomorphism between inner product spaces;

(3) T maps all the orthonormal bases in V to orthonormal bases in W ;

(4) T maps *one* orthonormal basis in V to a orthonormal basis in W.

Proof. It's clear that (1) $\implies$ (2) $\implies$ (3) $\implies$ (4), since T injective $\implies$ T is an isomorphism of vector space.

As for (4) $\implies$ (1), just expand everything using this orthonormal basis. □

Corollary 0.1.6

Inner product spaces with same dimensions are always isomorphic as inner product spaces.

Recall the positive definite matrices we defined earlier, we can also define *positive definite maps*: Let T be a *self-adjoint map*, if

$$\forall \alpha \in V \backslash \{0\}, \quad \langle T\alpha, \alpha \rangle > 0,$$

then we say T is positive definite.

The reason why we require T self-adjoint is that,

$$\langle T\alpha, \alpha \rangle = \langle \alpha, T\alpha \rangle = \overline{\langle T\alpha, \alpha \rangle} \implies \langle T\alpha, \alpha \rangle \in \mathbb{R}.$$

so we can talk about "positive" safely.

§0.2 Orthogonal maps and Unitary maps

Definition 0.2.1 (Orthogonal maps)**.** Let V be a real inner product space, the automorphisms of V (as inner product space) are called **orthogonal maps**, denoted the set by $\mathrm{O}(V)$.

When V is a complex inner product space, we use **unitary maps** and $\mathrm{U}(V)$ instead.

Proposition 0.2.2

Let V be an inner product space,

$$T \in \mathrm{O}(V) \iff T^* = T^{-1}.$$

Proof.

$$T \in \mathrm{O}(V) \iff \langle \alpha, \beta \rangle = \langle T\alpha, T\beta \rangle = \langle \alpha, T^*T\beta \rangle, \quad \forall \alpha, \beta \in V.$$

This also holds for $\mathrm{U}(V)$. □

Proposition 0.2.3

Let $A \in \mathbb{R}^{n\times n}$, TFAE:

- $A^tA = I_n$;
- The column (row) vectors of A form an orthonormal basis of $\mathbb{R}^n$.

Proof. Since A maps the standard basis to the column vectors of A, so the conclusion follows immediately (use A^t to get the row vectors). $\square$

Let $\mathrm{O}(n) = \{A \in \mathbb{R}^{n\times n} \mid A^tA = I_n\}$, and $\mathrm{U}(n) = \{A \in \mathbb{C}^{n\times n} \mid A^*A = I_n\}$. We can see that $A^tA = I_n \implies \det(A) = \pm 1$, and $A^*A = I_n \implies |\det(A)| = 1$.

Let $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det A = 1\}$, and $\mathrm{SU}(n) = \{A \in \mathrm{U}(n) \mid \det A = 1\}$. In the language of groups, $\mathrm{SO}(n)$ has only 2 coset in $\mathrm{O}(n)$, while the structure of the cosets of $\mathrm{SU}(n)$ in $\mathrm{U}(n)$ look like S^1.

Example 0.2.4

Let's look at some low dimensional orthogonal groups. $\mathrm{O}(1) = \{1, -1\}$, $\mathrm{SO}(1) = \{1\} = \mathrm{SU}(1)$, $\mathrm{U}(1) = \{z \mid |z| = 1\}$.

The group $\mathrm{SO}(2)$ is the rotations of $\mathbb{R}^2$:

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mid \theta \in [0, 2\pi) \right\}.$$

While $\mathrm{O}(2)$ also consisting of reflections.

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} \mid z, w \in \mathbb{C}, |z|^2 + |w|^2 = 1 \right\}.$$

In fact these groups are *lie groups*, which means they have the structure of differential manifolds. It's clear that $\mathrm{U}(1) \simeq \mathrm{SO}(2) \simeq S^1$, and we can see $\mathrm{SU}(2) \simeq S^3$.

Theorem 0.2.5 (QR-decomposition)

Any invertible matrix A can be uniquely decomposed to $Q \cdot R$, where $Q \in \mathrm{O}(n)$, R is an upper triangular matrix with positive diagonal entries. When $F = \mathbb{C}$, $\mathrm{O}(n)$ is replaced by $\mathrm{U}(n)$.

Proof. This is essentially Schmidt orthogonalozation. $\square$

Corollary 0.2.6 (Ivasawa decomposition, KAN decomposition)

For all $A \in \mathrm{GL}_n(\mathbb{R})$, it has a unique decomposition $A = A_kA_aA_n$, $A_k \in \mathrm{O}(n)$, A_a is diagonal, A_n is upper triangular matrix with diagonal entries 1. It also holds for $\mathbb{C}$.

Let $\mathcal{B}, \mathcal{B}'$ be orthonormal bases of V, $T \in L(V)$. We know that $[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P$ for some $P \in \mathrm{GL}(V)$. By orthogonality, P must be an orthogonal matrix, wich means $P^t = P^{-1}$.

Definition 0.2.7. Let $A, B \in \mathbb{R}^{n\times n}$, we say they are **orthogonally similar** if $A = P^{-1}BP$ for some $P \in \mathrm{O}(n)$. The name is changed to **unitarily similar** for complex matrices.

Theorem 0.2.8 (Schur triangularization theorem)
Let $F = \mathbb{C}, T \in L(V)$. There exists an orthonormal basis $\mathcal{B}$, such that $[T]_{\mathcal{B}}$ is upper triangular.

Proof. Recall that T is triangulable (which is always true in $\mathbb{C}$) iff there exists a T-invariant flag $\{0\} = W_0 \subset W_1 \subset \cdots \subset W_n = V$. We can take an orthonormal basis s.t. $W_k = \text{span}\{\alpha_1, \ldots, \alpha_k\}$.

Obviously T is upper triangular under this basis. $\square$