Linear Algebra II

Felix Chen

Contents

§1 Inner product spaces

In this section we always assume the base field to be $\mathbb{R}$ or $\mathbb{C}$.

§1.1 Inner product

Definition 1.1.1 (Inner product)**.** Let V be a vector space, an **inner product** on V is a function $\langle\cdot,\cdot\rangle : V \times V \to F$, $(\alpha,\beta) \mapsto \langle\alpha,\beta\rangle$ such that:

- $\langle\alpha+\beta,\gamma\rangle = \langle\alpha,\gamma\rangle + \langle\beta,\gamma\rangle$, $\langle c\alpha,\beta\rangle = c\langle\alpha,\beta\rangle$, i.e. the linearity of the first entry.
- $\langle\alpha,\beta\rangle = \overline{\langle\beta,\alpha\rangle}$. This implies the *conjugate linearity* of the second entry.
- $\alpha \neq 0 \implies \langle\alpha,\alpha\rangle > 0$.

Remark 1.1.2 — The reason why we require the conjugate property is that we want to make the inner product positive definite: otherwise $\langle i\alpha, i\alpha\rangle = i^2\langle\alpha,\alpha\rangle$.

The finite dimensional real inner product space is called **Euclid space**, and finite dimensional complex inner product space is called **unitary space**.

In fact the definition of inner space is related to the order in real numbers, so this is not a pure algebraic structure.

Example 1.1.3

Let $V = F^{n\times 1}$. Let $\alpha = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \beta = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, define $\langle\alpha,\beta\rangle = \sum_{j=1}^n x_j\overline{y_j} = \alpha^t\overline{\beta}$ to be the **standard inner product**.

Denote $\beta^* = \overline{\beta^t}$, then $\langle\alpha,\beta\rangle = \beta^*\alpha$.

Similarly when $V = F^{m\times n}$, $\langle A,B\rangle = \sum_{j,k} A_{jk}\overline{B_jk} = \mathrm{tr}(B^*A) = \mathrm{tr}(AB^*)$.

Definition 1.1.4 (Hermite matrices)**.** Let $A \in F^{n\times n}$, we say A is **Hermite** if $A^* = A$, and **anti-Hermite** if $A^* = -A$.

When $F = \mathbb{R}$, Hermite matrices are symmetrical matrices.

If we also have $\forall X \in F^{n\times 1}\backslash\{0\}$, $X^*AX > 0$, then we say A is **positive definite**.

Example 1.1.5

For all $Q \in \mathrm{GL}_n(F)$, $A = Q^*Q$ is positive definite.

Proposition 1.1.6

Let V be an n dimensional vector space, let $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ be a basis. For $\alpha, \beta \in V$, let $X = [\alpha]_{\mathcal{B}}$, $Y = [\beta]_{\mathcal{B}}$.

- If $A \in F^{n\times n}$ is positive definite, then

  $$\langle \alpha, \beta \rangle = Y^*AX = \sum_{j,k=1}^{n} A_{kj} x_j \overline{y_k}$$

  is an inner product.

- For any inner product $\langle \cdot, \cdot \rangle$, there exists a unique positive definite matrix A such that the above relations holds.

Proof. It's clear that Y^*AX is an inner product. (just check the definition)

For the latter part, let $A_{kj} = \langle \alpha_j, \alpha_k \rangle$, so A must be unique. By the conjugate linearity of inner product, so A constructed above indeed satisfies desired condition:

$$\langle \alpha, \beta \rangle = \left\langle \sum_{j=1}^{n} x_j \alpha_j, \sum_{k=1}^{n} y_k \alpha_k \right\rangle = \sum_{j,k=1}^{n} x_j \overline{y_k} \langle \alpha_j, \alpha_k \rangle$$

□

Let $T : V \to W$ be an injective linear map, and $\langle \cdot, \cdot \rangle_0$ is an inner product on W. Then T induces an inner product on V:

$$\langle \alpha, \beta \rangle = \langle T\alpha, T\beta \rangle_0, \quad \alpha, \beta \in V.$$

Since T injective, so T actually realizes V as a subspace of W, this inner product is just the original one restricted on the subspace.

Example 1.1.7

Let $V = W = F^{n\times 1}$, $\langle \cdot, \cdot \rangle_0$ is the standard inner product, $Q \in \mathrm{GL}_n(F)$. Then

$$\langle \alpha, \beta \rangle = \langle Q\alpha, Q\beta \rangle_0 = \beta^*(Q^*Q)\alpha.$$

With an inner product, we can assign a "length" to each vector: $\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle}$. It's clear that:

$$\|c\alpha\| = |c|\|\alpha\|, \quad \|\alpha\| > 0, \forall \alpha \neq 0.$$

Proposition 1.1.8 (Polarization identity)
When $F = \mathbb{R}$,

$$\langle \alpha, \beta \rangle = \frac{1}{4}\left(\|\alpha+\beta\|^2 - \|\alpha-\beta\|^2\right).$$

When $F = \mathbb{C}$,

$$\langle \alpha, \beta \rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|\alpha + i^k\beta\|^2.$$

Remark 1.1.9 — This means, *inner product is totally determined by length function.*

Proposition 1.1.10 (Cauchy-Schwarz inequality)

$$|\langle \alpha, \beta \rangle| \le \|\alpha\|\|\beta\|.$$

The equality holds iff α, β linearly dependent.

Proof. WLOG $\alpha, \beta \neq 0$. Let $\gamma = \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2}\alpha$ be the orthogonal projection of β on $\alpha^\perp$.

We can check that $\langle \alpha, \gamma \rangle = 0$, so

$$0 \le \|\gamma\|^2 = \langle \gamma, \beta \rangle = \|\beta\|^2 - \frac{\langle \alpha, \beta \rangle^2}{\|\alpha\|^2},$$

which gives the desired inequality, equality iff $\gamma = 0$ iff α, β linearly dependent. □

Proposition 1.1.11 (Triangle inequality)

$$\|\alpha + \beta\| \le \|\alpha\| + \|\beta\|.$$

Proof. Square both sides and use Cauchy-Schwarz. □

This means our "length" function is in fact a **norm**.

§1.2 Orthogonality

Definition 1.2.1 (Orthogonality)**.** Let $\alpha, \beta \in V$, we say $\alpha \perp \beta$ if $\langle \alpha, \beta \rangle = 0$.

We can introduce "angles" as well:

Definition 1.2.2 (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 1.2.3 (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 1.2.4

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

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

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

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

Proposition 1.2.5

If $S = \{\alpha_1, \dots, \alpha_m\}$ 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 1.2.6 (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 1.2.7 — 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 1.2.8
Let $S \subset V \backslash \{0\}$ be orthogonal(-normal), then S can be extended to an orthogonal(-normal) basis.

Proposition 1.2.9
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 1.2.10 (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 1.2.11
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. $\phi \in \mathrm{Hom}_{\mathbb{R}}(V, V^*)$. 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 1.2.12 (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.

§1.3 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.

$$\begin{tikzcd} V^* \arrow[r, "T^t"] & V^* \\ V \arrow[u, "\phi"] \arrow[r, "T^*"] & V \arrow[u, "\phi"] \end{tikzcd}$$

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 1.3.1

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 1.3.2

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

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 1.3.3

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 $[p_W]_{\mathcal{B}} = \operatorname{diag}\{I_k, 0\}$, 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 1.3.4

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 1.3.5

Let V, W be finite dimensional inner product spaces, $\dim V = \dim W$, $T \in \operatorname{Hom}(V, W)$, 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 1.3.6

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

Definition 1.3.7 (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 1.3.8

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 1.3.9

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

- $A^t A = I_n$;
- 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). □

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