Linear Algebra II

Felix Chen

Contents

§1 Bilinear forms

Let V be a finite dimensional vector space, $\dim V = n$.

Definition 1.0.1. Let $F = \mathbb{C}$, we say a function $f : V \times V \to V$ is a **semi bilinear form** if:

- $f(c_1\alpha + \beta, \gamma) = c_1 f(\alpha, \gamma) + f(\beta, \gamma)$;
- $f(\alpha, c_1\beta + \gamma) = \overline{c}_1 f(\alpha, \beta) + f(\alpha, \gamma)$.

Let $\mathrm{Form}(V)$ denote the (semi) bilinear forms on (complex) real vector space V.

For $f \in \mathrm{Form}(V)$, fix a basis $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ of V, let $[f]_{\mathcal{B}} \in F^{n\times n}$ be the matrix

$$([f]_{\mathcal{B}})_{jk} = f(\alpha_k, \alpha_j).$$

which is called **the matrix of f under $\mathcal{B}$**.

For $\alpha = \sum_{k=1}^{n} x_k\alpha_k, \beta = \sum_{j=1}^{n} y_j\alpha_j \in V$. It's clear that

$$f(\alpha, \beta) = \sum_{j,k=1}^{n} x_k\overline{y}_j f(\alpha_k, \alpha_j) = \sum_{j,k=1}^{n} x_k\overline{y}_j ([f]_{\mathcal{B}})_{jk} = [\beta]_{\mathcal{B}}^{*}[f]_{\mathcal{B}}[\alpha]_{\mathcal{B}}.$$

From this we know that the map $\mathrm{Form}(V) \to F^{n\times n}, f \mapsto [f]_{\mathcal{B}}$ is a linear isomorphism. Since if $[f]_{\mathcal{B}} = 0$, then $f(\alpha, \beta) = 0$ for all $\alpha, \beta \in V$. Thus it's injective. Obviously it's surjective and linear, so

$$\dim \mathrm{Form}(V) = n^2$$

Example 1.0.2

Let $A \in F^{n\times n}$. Let $f \in \mathrm{Form}(F^{n\times 1})$ be

$$f(X, Y) = Y^{*}AX, \quad \forall X, Y \in F^{n\times 1}.$$

Let $\mathcal{B}$ be the standard basis of F, it's clear that $[f]_{\mathcal{B}} = A$.

Proposition 1.0.3
Let $\mathcal{B}' = \{\alpha_1', \dots, \alpha_n'\}$ be another basis of V, $P \in \mathrm{GL}_n(F)$ satisfies

$$(\alpha_1', \dots, \alpha_n') = (\alpha_1, \dots, \alpha_n)P.$$

Then $[f]_{\mathcal{B}'} = P^*[f]_{\mathcal{B}}P$.

Proof. Since $[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$, just plug this into the definition of $[f]_{\mathcal{B}}$, the rest is trivial. □

Definition 1.0.4. Let $f \in \mathrm{Form}(V)$.

- When $F = \mathbb{R}$, if $\forall \alpha, \beta \in V$ we have $f(\alpha, \beta) = f(\beta, \alpha)$, then we say f is symmetrical (also called Hermite);
- When $F = \mathbb{C}$, if $\forall \alpha, \beta \in V$ we have $f(\alpha, \beta) = \overline{f(\beta, \alpha)}$, we say f is Hermite.

Proposition 1.0.5
When $F = \mathbb{C}$, f Hermite $\iff f(\alpha, \alpha) \in \mathbb{R}, \forall \alpha \in V$.

Proof. For the " $\Longleftarrow$ " direction, consider $f(\alpha+\beta, \alpha+\beta) \in \mathbb{R}$. Expanding we'll get $f(\alpha, \beta) + f(\beta, \alpha) \in \mathbb{R}$, i.e.

$$f(\alpha, \beta) + f(\beta, \alpha) = \overline{f(\alpha, \beta)} + \overline{f(\beta, \alpha)}.$$

Replace α with $i\alpha$, we get

$$f(\alpha, \beta) - f(\beta, \alpha) = -\overline{f(\alpha, \beta)} + \overline{f(\beta, \alpha)}.$$

Combining two equations we get the conclusion. □

Definition 1.0.6. Let $f \in \mathrm{Form}(V)$ be an Hermite form. If $\forall \alpha \in V \backslash \{0\}$, $f(\alpha, \alpha) > 0$, we say f is **positive definite**.

Similarly we can define negative definite and semi positive definite.

Note that a positive definite Hermite form is nothing but an inner product.

§1.1 Positive define matrices

In this section we'll dig deeper into properties of positive definite matrices.

It's clear that if a matrix A is positive definite, then A is inversible, and P^*AP is also positive definite. In particular, P^*P is positive definite.

Theorem 1.1.1 (Cholesky decomposition)
Let $A \in F^{n\times n}$ be a positive definite matrix, there exists a unique upper triangular matrix R with positive diagonal entries s.t. $A = R^*R$.

Proof. Consider the inner product $f(X, Y) = Y^*AX$. Let the standard inner product on V be $f_0(X, Y) = Y^*X$.

Since inner product spaces with same dimensions are isomorphic, so there exists a matrix $R \in \mathrm{GL}_n(F)$, such that

$$R : (F^{n\times 1}, f) \to (F^{n\times 1}, f_0), \quad X \mapsto RX$$

is an isomorphism of inner product space, i.e. $f_0(RX, RY) = f(X, Y)$. This is equivalent to $A = R^*R$.

For any $P \in GL_n(F)$, P is also an isomorphism of $(F^{n\times 1}, f) \to (F^{n\times 1}, f_0)$ iff RP^{-1} preserves the inner product f_0, iff $RP^{-1} \in \mathrm{O}(n)$ or $\mathrm{U}(n)$.

By QR decomposition, $R = RP^{-1} \cdot P$, so there must be a unique P s.t. P upper triangular with positive diagonal entries. □

Corollary 1.1.2

A positive definite $\implies \det A > 0$.

Definition 1.1.3. Let $A \in F^{n\times n}$, for $1 \le k \le n$, define

$$\Delta_k(A) := \det(A^{1\le j\le k}_{1\le i\le k})$$

be the **leading principal minor**.

Theorem 1.1.4

Let $A \in F^{n\times n}$ be an Hermite matrix. Then A positive definite $\iff \Delta_k(A) > 0$, $k = 1, \ldots, n$.

Lemma 1.1.5 (LU decomposition)

Let F be *any field*. For $A \in \mathrm{GL}_n(F)$, the followings are equivalent:

- $\Delta_k(A) \ne 0, k = 1, \ldots, n$;
- $A = LU$, where L lower triangular, and U upper triangular with diagonal entries 1.

Proof. On one hand, Let L_k, U_k be the top-left $k \times k$ submatrix of L, U, since L, U inversible, L_k, U_k inversible. By the triangular property, $\Delta_k(A) = \det(L_k U_k) \ne 0$.

On the other hand, it's sufficient to prove:

$$\exists N \text{ strictly upper triangular}, A(N + I_n) \text{ lower triangular}$$

Let A_k be the k-th leading principal submatrix of A, and $\alpha_{k+1}, \beta_{k+1} \in F^{n\times 1}$ the $(k+1)$-th column of A, N.

Now compute the first k rows of the $(k+1)$-th column of $A(N+I)$, which is equal to $A_k\beta'_{k+1} + \alpha'_{k+1}$, where $\alpha'_{k+1}, \beta'_{k+1}$ is the first k entries of $\alpha_{k+1}, \beta_{k+1}$.

Since A_k inversible, $\exists \beta'_{k+1}$ s.t. $A_k\beta'_{k+1} + \alpha'_{k+1} = 0$.

Hence these β'_{k+1} forms a strictly upper triangular matrix N, as desired. □

Proof of the theorem. Let A be an Hermite matrix, if A positive definite, then $\det A \ge 0$.

Let A_k be the upper left $k \times k$ submatrix of A. For $X \in F^{k\times 1}\backslash\{0\}$, we have

$$X^* A_k X = \begin{pmatrix} X \\ 0 \end{pmatrix}^* A \begin{pmatrix} X \\ 0 \end{pmatrix} > 0.$$

Hence A_k positive definite, $\det A_k = \Delta_k(A) \ge 0$.

Conversely, by our lemma let $A = LU$, let $D = (U^*)^{-1}L$, $A = U^*DU$.
Hence A Hermite $\implies$ D Hermite. Moreover D is lower triangular, so D is diagonal.
Some computation yields that $A_k = U_k^* D_k U_k$. Therefore

$$\Delta_k(A) \geq 0 \implies \det(U_k^* D_k U_k) \geq 0 \implies \det D_k \geq 0.$$

From this we deduce that all the diagonal entries of D are positive, so D positive definite $\implies$ A positive definite. □

§1.2 Bilinear forms on inner product spaces

Let V be an inner product space, given a basis of V, recall that there are two linear isomorphism:

$$\mathrm{Form}(V) \to F^{n\times n}, f \mapsto [f]_{\mathcal{B}} \quad L(V) \to F^{n\times n}, T \mapsto [T]_{\mathcal{B}}$$

Hence we can define a map $\mathrm{Form}(V) \to L(V)$ by composing these two isomorphism. Denote this map by $f \mapsto T_f$. It seems like this map also depends on the choice of the basis, but in fact it's independent as long as $\mathcal{B}$ is orthonormal!

Let $\mathcal{B}'$ be another orthonormal basis, then $[T_f]_{\mathcal{B}'} = P^{-1}[T_f]_{\mathcal{B}}P$, while $[f]_{\mathcal{B}'} = P^*[f]_{\mathcal{B}}P$, but P is orthogonal (or unitary), so $P^{-1} = P^*$, i.e. T_f doesn't change under the new basis.

Since T_f do not depend on the basis, thus we wonder whether we can define this map intrinsically.

Proposition 1.2.1

For all $T \in L(V)$, T induces a (semi) bilinear form $f(\alpha, \beta) = \langle T\alpha, \beta\rangle$. We claim that this map $\mathcal{F}$ gives an isomorphism of $L(V)$ and $\mathrm{Form}(V)$.

Proof. Clearly $\mathcal{F}$ is injective:

$$\langle T\alpha, \beta\rangle = 0, \forall \beta \implies T\alpha = 0,$$

thus $\ker \mathcal{F} = \{0\}$.

By dimenional reasons $\mathcal{F}$ must be an isomorphism. □

By considering $\mathcal{F}^{-1}$, we get an one-to-one map $f \mapsto T_f$ s.t.

$$f(\alpha, \beta) = \langle T_f\alpha, \beta\rangle.$$

We'll see that this definition coincide with the initial one. In fact it's sufficient to prove $[T_f]_{\mathcal{B}} = [f]_{\mathcal{B}}$, which is just a bunch of computation ;)

Remark 1.2.2 — We can construct T_f explicitly from f:
The inner product gives a conjugate linear isomorphism

$$\Phi : V \to V^*, \quad \Phi(\alpha)(\beta) = \langle \beta, \alpha\rangle = \overline{\langle \alpha, \beta\rangle}.$$

Similarly, $f \in \mathrm{Form}(V)$ gives a conjugate linear map

$$\Phi_f : V \to V^*, \quad \Phi_f(\alpha)(\beta) = \overline{f(\alpha, \beta)}.$$

Then $T = \Phi^{-1} \circ \Phi_f$ is the desired linear map:

$$\langle T\alpha, \beta \rangle = \overline{\Phi(T\alpha)(\beta)} = \overline{\Phi_f(\alpha)(\beta)} = f(\alpha, \beta).$$

Hence all the properties of linear maps can be carried over to the forms, and vice versa (using the matrix representation).

Corollary 1.2.3
Let $F = \mathbb{C}$, $T \in L(V)$, T self-adjoint iff $\langle T\alpha, \alpha \rangle \in \mathbb{R}, \forall \alpha \in V$.

Proof. T self-adjoint iff f Hermite iff $f(\alpha, \alpha) \in \mathbb{R}$. □

Corollary 1.2.4
Let $f \in \mathrm{Form}(V)$.

- If f Hermite, there exists an orthonormal basis of V s.t. $[f]_{\mathcal{B}}$ is real diagonal.
- If $F = \mathbb{C}$, there exists an orthonormal basis such that $[f]_{\mathcal{B}}$ upper triangular.

§1.3 Spectral decomposition

Theorem 1.3.1 (Spectral decomposition of normal maps)
Let $T \in L(V)$ be a self-adjoint map (or normal map in complex field), let $\sigma(T) = \{c_1, \ldots, c_k\}$, $P_i \in L(V)$ are the projection onto V_{c_i}. Then for any $f \in F[x]$, we have

$$f(T) = \sum_{i=1}^{k} f(c_i) P_i.$$

In particular, $T = \sum_{i=1}^{k} c_i P_i$.

Proof. Consider the orthogonal direct sum

$$V = \bigoplus_{i=1}^{k} V_{c_i},$$

since previously we've proven that T is orthogonally diagonalizable (or unitarily diagonalizable).

Using this decomposition, the conclusion is somewhat trivial. □

Corollary 1.3.2
Each P_i is a polynomial of T.

Proof. Take $f_i \in F[x]$ s.t. $f_i(c_i) = \delta_{ij}$. Then $f_i(T) = \sum_{j=1}^{k} f_i(c_j) P_j = P_i$. □

Using simialr technique, we can consider functions other than polynomials of T, defined by $\phi(T) = \sum_{i=1}^k \phi(c_i)T$. By Lagrange interpolation, we can always find a polynomial p s.t. $p(c_i) = \phi(c_i)$ for all $c_i \in \sigma(T)$.

Example 1.3.3

If T semi positive definite normal matrix, $\sigma(T) \subset [0, +\infty)$, so we can define $\sqrt{T} = \sum_{i=1}^k \sqrt{c_i} P_i$.

Proposition 1.3.4

T self-adjoint (normal) $\implies \phi(T)$ self-adjoint (normal); $\sigma(\phi(T)) = \phi(\sigma(T))$.

Proof. Note that T and $\phi(T)$ are diagonal matrices under orthonormal basis of V_{c_i}. □

Theorem 1.3.5

Let $T \in L(V)$ be semi positive definite.

- $\sqrt{T}$ semi positive definite, and $\sqrt{T}^2 = T$.
- T positive definite $\iff \sqrt{T}$ positive definite.
- If $S \in L(V)$ semi positive definite and $S^2 = T$, then $S = \sqrt{T}$.

Proof. Since $[\sqrt{T}]_{\mathcal{B}} = \operatorname{diag}(\sqrt{c_1} I_{d_1}, \ldots, \sqrt{c_k} I_{d_k})$, the first two statements are trivial.

Let $\sigma(S) = \{s_1, \ldots, s_r\}$, $V_i = \ker(S - s_i \operatorname{id})$. Since S self-adjoint, $V = \bigoplus_{i=1}^r V_i$.

For any $\alpha \in V_i, T\alpha = S^2\alpha = s_i^2\alpha$, thus $V_i \subset \ker(T - s_i^2 \operatorname{id})$. Since $s_i \geq 0$, $\sqrt{T} = S$. □

Note that T^*T is always positive definite, so we can consider $\sqrt{T^*T}$. We call the eigen-values of $\sqrt{T^*T}$ **singular values** of T.

In some sense, $\sqrt{T^*T}$ is a semi positive approximation of T:

Lemma 1.3.6

For any $\alpha \in V$, $\|T\alpha\| = \|\sqrt{T^*T}\alpha\|$. In particular, $\ker T = \ker \sqrt{T^*T}$.

Proof. Let $N = \sqrt{T^*T}$,

$$\|N\alpha\|^2 = \langle N\alpha, N\alpha\rangle = \left\langle N^2\alpha, \alpha\right\rangle = \langle T^*T\alpha, \alpha\rangle = \langle T\alpha, T\alpha\rangle = \|T\alpha\|^2.$$

□

Theorem 1.3.7 (Polar decomposition)
Let $T \in L(V)$,

(1) There exists $U \in L(V)$ orthogonal or unitary, $N \in L(V)$ semi positive definite, $T = UN$.

(2) We must have $N = \sqrt{T^*T}$.

(3) T invertible iff N positive definite, in this case U is unique.

Remark 1.3.8 — This is a generalization of $z = re^{i\theta}$ in $\mathbb{C}$. That's where the name comes from.

Proof. If (1) holds, then

$$T^* = NU^* \implies T^*T = NU^*UN = N^2 \implies N = \sqrt{T^*T}.$$

Since T, N are semi positive definite, T invertible iff T positive definite. Now we must have $U = TN^{-1}$, which is unique.

To prove (1), when T invertible, let N, U as above, by our lemma,

$$\|U\alpha\| = \|TN^{-1}\alpha\| = \|\alpha\|$$

Thus U is orthogonal or unitary.

When T is not invertible, $\ker T = \ker N$, thus $\exists U_1 : \mathrm{Im}(N) \to \mathrm{Im}(T)$ s.t. $T = U_1 N$. (Just take $N\alpha \mapsto T\alpha$)

Moreover U_1 is an isomorphism of inner product space: $\|U_1 N\alpha\| = \|T\alpha\| = \|N\alpha\|$. So U_1 preserves inner product and hence injective. By dimension reasons, U_1 must be an isomorphism.

Now we can take an arbitary isomorphism $U_2 : \mathrm{Im}(N)^\perp \to \mathrm{Im}(T)^\perp$, $U := U_1 \oplus U_2$ is the desired map. $\square$

Looking back at the singular values, consider the image of unit sphere $S \subset V$ under T, $N(S)$ is an ellipsoid:

$$N(S) = \left\{ \sum_{i=1}^n c_i x_i \alpha_i : \sum_{i=1}^n x_i^2 = 1 \right\}.$$

Since $T = UN$, U is a rotation, so $T(S)$ is also an ellipsoid, whose axes lengths are $2c_i$, where c_i are singular values of T.

Corollary 1.3.9 (Singular value decomposition, SVD)
Let $A \in F^{n\times n}$, then there exists decomposition $A = U_1 D U_2$, where D is a diagonal matrix with non-negative entries, U_1, U_2 are orthogonal or unitary matrices.

Proof. Consider the polar decomposition $A = UN$, let $N = PDP^{-1}$, where P orthogonal or unitary, D non-negative diagonal. Thus we can take $U_1 = UP, U_2 = P^{-1}$.

Note that the diagonal entries of D is precisely the singular value of A. $\square$

Corollary 1.3.10

Let $T \in L(V)$, then T map *some* orthogonal basis to another orthogonal basis.

Proof. Let $T = UN$ be the polar decomposition. Let $\alpha_1, \dots, \alpha_n$ be an orthonormal basis s.t. N diagonal, then

$$T\alpha_i = UN\alpha_i = c_i U\alpha_i,$$

obviously $c_i U\alpha_i$ consititude an orthogonal basis. □

§1.4 Further on normal maps

For $\theta \in \mathbb{R}$, let Q_θ be the rotation of angle θ. The main goal of this section is to prove:

Theorem 1.4.1

Let V be a finite dimensional real inner product space, $T \in L(V)$ normal. There exists an orthonormal basis $\mathcal{B}$ s.t.

$$[T]_\mathcal{B} = \operatorname{diag}(a_1, \dots, a_l, r_1 Q_{\theta_1}, \dots, r_m Q_{\theta_m}),$$

where $a_i \in \mathbb{R}, r_j > 0, \theta_j \in (0, \pi)$.

Let's look at a corollary of this theorem first:

Corollary 1.4.2

If T orthogonal, then

$$[T]_\mathcal{B} = \operatorname{diag}(I_{l_1}, -I_{l_2}, Q_{\theta_1}, \dots, Q_{\theta_m}).$$

Proof. Applying the theorem, since each block is orthogonal, $a_i = \pm 1$, $r_j = 1$. □

This gives us a comprehension of rotations in higher dimensional spaces.

Here we'll present multiple proofs to emphasize some intermediate result.

Proposition 1.4.3

Let T be a normal map, if $W \subset V$ is T-invariant, then T_W is also normal.

Proof. First note that $W, W^\perp$ are T^*-invariant. For $\alpha, \beta \in W$, we have

$$\langle (T_W)^*\alpha, \beta \rangle = \langle \alpha, T_W\beta \rangle = \langle \alpha, T\beta \rangle = \langle T^*\alpha, \beta \rangle .$$

Thus $(T_W)^* = T^*_W$. The conclusion follows. □

Proposition 1.4.4

Let T be a normal map, there exists an orthogonal decomposition $V = \bigoplus_{i=1}^k V_i$, such that each V_i is T-invariant, and T_{V_i} simple.

Proof. Note that if W is T-invariant, then $W^\perp$ is also T-invariant. By induction and the previous proposition this is trivial. $\square$

Therefore to prove Theorem 1.4.1, we only need to prove the case when T is simple.

Proof of Theorem 1.4.1. WLOG $\dim V > 1$.

Since T simple $\implies f_T \in \mathbb{R}[x]$ prime, thus $\deg f_T = 2$, $\dim V = 2$ and $f_T = (x-c)(x-\overline{c})$.

Take any orthonormal basis $\mathcal{B} = \{\alpha_1, \alpha_2\}$, let $r = |c|$, $A = r^{-1}[T]_{\mathcal{B}}$. Clearly A normal and $\sigma(A) = \{r^{-1}c, r^{-1}\overline{c}\}$, so A is unitarily similar to $\mathrm{diag}(r^{-1}c, r^{-1}\overline{c})$, A is unitary.

Moreover A is a real matrix so A orthogonal, and $\det A = 1$, thus $A = Q_\theta, \theta \in [0, 2\pi]$.

At last by T has no eigenvector, and we can change α_2 to $-\alpha_2$, so we can require $\theta \in (0, \pi)$. $\square$

Proposition 1.4.5

Let $T \in L(V)$, then $\ker(T)^\perp = \mathrm{im}(T^*), \mathrm{im}(T)^\perp = \ker(T^*)$.

Proof. Trivial, just some computation. $\square$

Proposition 1.4.6

Let $T \in L(V)$, $\sigma(T^*) = \overline{\sigma(T)}$,

$$\forall c \in \sigma(T), \quad \dim\ker(T - cI) = \dim\ker(T^* - \overline{c}I).$$

Proof. By the previous proposition,

$$\dim\ker(T - cI) = n - \dim\mathrm{im}(T^* - \overline{c}I) = \dim\ker(T^* - \overline{c}I)$$

which also implies $\sigma(T) = \overline{\sigma(T^*)}$. $\square$

Proposition 1.4.7

If T normal, then $\ker(T - cI) = \ker(T^* - \overline{c}I)$.

Proof. Let $W = \ker(T - cI)$, T_W^* is just $(c\,\mathrm{id}_W)^* = \overline{c}\,\mathrm{id}_W$. Thus $W \subset \ker(T^* 0 \overline{c}I)$, by dimensional reasons they must be equal. $\square$

Proposition 1.4.8

Let T be a normal map, $f, g \in F[x]$ coprime $\implies \ker(f(T)) \perp \ker(g(T))$.

Proof. Since $g(T)^* = \overline{g}(T^*)$, $g(T)$ is normal, thus $\ker(g(T))^\perp = \mathrm{im}(g(T))$.

Let $W = \ker(f(T))$, let $a, b \in F[x]$ s.t. $af + bg = 1$, so $a(T)f(T) + b(T)g(T) = \mathrm{id}_V$. Restrict this equation to W, we get $b(T)_W g(T)_W = \mathrm{id}_W$, hence $W \subset \mathrm{im}(g(T))$. $\square$

Proposition 1.4.9

Let T be a normal map,

- The primary decomposition of T are orthogonal decomposition;
- The cyclic decomposition of T can be orthogonal.

Proof. The first one is trivial by previous proposition.

For cyclic decomposition, we proceed by induction on $\dim V$.

Let $\alpha_1 \in V$ s.t. $p_{\alpha_1} = p_r$, then $(R\alpha_1)^\perp$ are T-invariant, use induction hypo on it and we're done. □

Remark 1.4.10 — This means the primary cyclic decomposition of T can also be orthogonal.

This gives the second proof of Theorem 1.4.1:

Proof. WLOG T normal and primary cyclic, then p_T is primary, and T normal $\implies$ T semisimple, so p_T has no multiple factors, thus p_T prime, which proves the result. □

Next we present the third proof:

Proposition 1.4.11

If $A, B \in \mathbb{R}^{n\times n}$ are unitarily similar, then they are orthogonally similar.

Proposition 1.4.12 (QS decomposition)

For any unitary matrix U, $U = QS$ where Q real orthogonal, S unitary and symmetrical. Moreover $\exists f \in \mathbb{C}[x]$ s.t. $S = f(U^tU)$.