Linear Algebra II

Felix Chen

Contents

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

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, $\det A_k = \Delta_k(A) \geq 0$.

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. □

§0.1 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 0.1.1

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 0.1.2 — 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 0.1.3

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 0.1.4

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.

§0.2 Spectral decomposition

Theorem 0.2.1 (Spectral decomposition of normal maps)
Let $T \in L(V)$ be a self-adjoint map (or normal map in complex field), let $\sigma(T) = \{c_1, \dots, c_k\}$, $P_i \in L(V)$ 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 0.2.2
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 0.2.3
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 0.2.4
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 0.2.5
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 $[\sqrt{T}]_{\mathcal{B}} = \operatorname{diag}(\sqrt{c_1} I_{d_1}, \dots, \sqrt{c_k} I_{d_k})$, the first two statements are trivial.

Let $\sigma(S) = \{s_1, \dots, s_r\}$, $V_i = \ker(S - s_i \operatorname{id})$. 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 0.2.6

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 0.2.7 (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 0.2.8 — 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 $\exists U_1 : \operatorname{Im}(N) \to \operatorname{Im}(T)$ s.t. $T = U_1N$. (Just take $N\alpha \mapsto T\alpha$)

Moreover U_1 is an isomorphism of inner product space: $\|U_1N\alpha\| = \|T\alpha\| = \|N\alpha\|$. 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 $U_2 : \operatorname{Im}(N)^\perp \to \operatorname{Im}(T)^\perp$, $U := U_1 \oplus U_2$ is the desired map. ☐

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 0.2.9 (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. □

Corollary 0.2.10

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. □

§0.3 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 0.3.1

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 0.3.2

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 0.3.3

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 0.3.4

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. □

Therefore to prove Theorem 0.3.1, we only need to prove the case when T is simple.

Proof of Theorem 0.3.1. WLOG $\dim V > 1$.

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

Take any orthonormal basis $\mathcal{B} = \{\alpha_1, \alpha_2\}$, let $r = |c|$, $A = r^{-1}[T]_{\mathcal{B}}$. Clearly A normal and $\sigma(A) = \{r^{-1}c, r^{-1}\bar{c}\}$, so A is unitarily similar to $\operatorname{diag}(r^{-1}c, r^{-1}\bar{c})$, 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 $\theta \in (0, \pi)$. □

Proposition 0.3.5

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

Proof. Trivial, just some computation. □

Proposition 0.3.6

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

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

Proof. By the previous proposition,

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

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

Proposition 0.3.7

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

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

Proposition 0.3.8

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 $W \subset \mathrm{im}(g(T))$. □

Proposition 0.3.9

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

This gives the second proof of Theorem 0.3.1:

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 0.3.11

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

Lemma 0.3.12 (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. $S = f(U^t U)$.

Proof. Let $\sigma(U^tU) = \{c_1, \dots, c_k\}$. We can take a polynomial $f \in \mathbb{C}[x]$ s.t. $f(c_i)^2 = c_i$.

Since U is unitary, $|c_i| = 1 \implies |f(c_i)| = 1$.

Let $S = f(U^tU)$, we claim that S unitary and $S^2 = U^tU$.

Let $U^tU = P\operatorname{diag}(c_1, \dots, c_k)P^{-1}$, where P is unitary, then $S = P\operatorname{diag}(f(c_1), \dots, f(c_k))P^{-1}$ is unitary, and clearly $S^2 = U^tU$.

Let $Q = US^{-1}$, then Q unitary. Since S symmetrical, $S^{-1} = S^* \implies \overline{S^{-1}} = S^t = S$,

$$\overline{Q}Q^{-1} = \overline{U}SSU^{-1} = \overline{U}U^tUU^{-1} = I_n.$$

Hence $\overline{Q} = Q$, Q is real orthogonal. □

Return to the original proposition. Let A, B be real matrices unitarily similar, let $B = UAU^{-1}$, taking the conjuate we get

$$UAU^{-1} = \overline{U}AU^t \implies U^tUA = AU^tU.$$

Let $U = QS$, then $AS = SA$. We have

$$B = UAU^{-1} = QSAS^{-1}Q^{-1} = QAQ^{-1}.$$

Therefore A, B are orthogonally similar. □

Corollary 0.3.13

Let A, B be normal matrices, TFAE:

(1) A, B are unitarily similar (or orthogonally similar);

(2) A, B are similar;

(3) $f_A = f_B$.

Proof. We only need to prove (3) $\implies$ (1).

When $F = \mathbb{C}$, A, B are unitarily similar to diagonal matrices D_1, D_2. Since $f_A = f_B$, D_1, D_2 only differ by a permutation, hence unitarily similar.

When $F = \mathbb{R}$, by the previous proposition and proof for $\mathbb{C}$, we get the result. □

The third proof of Theorem 0.3.1 is to factorize $f_T \in \mathbb{R}[x]$ and use the above corollary.

At last we prove another property of normal maps:

Proposition 0.3.14

Let A be a normal matrix, then A^* is a complex polynomial of A.

Proof. Use the spectral decomposition. □

§1 Bilinear forms

In this section we study the bilinear forms on generic fields. Let $M^2(V)$ denote all the bilinear forms on V.

For $f \in M^2(V)$, Let $(f(\alpha_i, \alpha_j))_{ij}$ be the matrix of f under basis $\{\alpha_i\}$. (Note that this differs by a transpose with previous section)

Obviously $M^2(V) \to F^{n\times n}$ by $f \mapsto [f]_{\mathcal{B}}$ is a linear isomorphism.

Proposition 1.0.1

Let $\mathcal{B}, \mathcal{B}'$ be two basis, P is the transformation matrix between them, for all $f \in M^2(V)$ we have $[f]_{\mathcal{B}'} = P^t [f]_{\mathcal{B}} P$.

Proof. Trivial. □

If $A = P^t BP$ for some $P \in \mathrm{GL}(V)$, we say A, B are **congruent**.
A bilinear form will induce two linear maps $V \to V^*$, namely L_f, R_f:

$$L_f(\alpha)(\beta) = R_f(\beta)(\alpha) = f(\alpha, \beta).$$

Proposition 1.0.2

For any basis $\mathcal{B}$, we have $\operatorname{rank} L_f = \operatorname{rank} R_f = \operatorname{rank}[f]_{\mathcal{B}}$. This number is called the rank of f, denoted by $\operatorname{rank} f$.

If $\operatorname{rank} f = n$, we say f is non-degenrate, this is equivalent to L_f invertible or R_f invertible.

§1.1 Some special bilinear forms

Definition 1.1.1. For $f \in M^2(V)$,

- If $f(\alpha, \beta) = f(\beta, \alpha), \forall \alpha, \beta \in V$, then we say f is **symmetrical**.
- If $f(\alpha, \beta) = -f(\beta, \alpha), \forall \alpha, \beta \in V$, we say f is **anti-symmetrical**.
- If $f(\alpha, \alpha) = 0, \forall \alpha \in V$, we say f is **alternating**.

We denote the above functions by $S^2(V), A^2(V), \Lambda^2(V)$.

We can see that $\Lambda^2(V) \subset A^2(V)$, and they are all subspaces of $M^2(V)$.

Proposition 1.1.2

If $\operatorname{char} F \neq 2$, then $A^2(V) = \Lambda^2(V)$, and $M^2(V) = A^2(V) \oplus S^2(V)$.

Proof. Already proved in last semester. □

Proposition 1.1.3

Let $\mathcal{B}$ be any basis of V,

- f symmetrical $\iff$ $[f]_{\mathcal{B}}$ symmetrical;
- f anti-symmetrical $\iff$ $[f]_{\mathcal{B}}$ anti-symmetrical;
- f alternating $\iff$ $[f]_{\mathcal{B}}$ anti-symmetrical and the diagonal entries are all zero.