Linear Algebra II

Felix Chen

Contents

Theorem 0.0.1

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 0.0.2

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 0.0.3 (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 0.0.4 — 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. $T = U_1N$. (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 $U_2 : \mathrm{Im}(N)^\perp \to \mathrm{Im}(T)^\perp$, $U := U_1 \oplus U_2$ 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 0.0.5 (Singular value decomposition, SVD)
Let $A \in F^{n\times n}$, then there exists decomposition $A = U_1DU_2$, 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 0.0.6
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. □

§0.1 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 0.1.1
Let V be a finite dimensional real inner product space, $T \in L(V)$ normal. There exists an orthonormal basis $\mathcal{B}$ s.t.

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

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 0.1.2
If T orthogonal, then

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

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 0.1.3
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 0.1.4
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 0.1.1, we only need to prove the case when T is simple.

Proof of Theorem 0.1.1. 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 $\theta \in (0, \pi)$. $\square$

Proposition 0.1.5

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

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

Proposition 0.1.6

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

Proposition 0.1.7

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. $\square$

Proposition 0.1.8

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

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

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 $W \subset \operatorname{im}(g(T))$. $\square$

Proposition 0.1.9

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. $\square$

Remark 0.1.10 — This means the primary cyclic decomposition of T can also be orthogonal.

This gives the second proof of Theorem 0.1.1:

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. $\square$

Next we present the third proof:

Proposition 0.1.11

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

Lemma 0.1.12 (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. $S = f(U^tU)$.

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. $\square$

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. $\square$

Corollary 0.1.13

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 0.1.1 is to factorize $f_T \in \mathbb{R}[x]$ and use the above corollary.

At last we prove another property of normal maps:

Proposition 0.1.14

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 $A = P^tBP$ 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.

Definition 1.1.4 (Quadratic forms)**.** Let $q : V \to F$ be a function, we say q is a **quadratic form** if there exists $f \in M^2(V)$ s.t.

$$q(\alpha) = f(\alpha,\alpha), \quad \forall \alpha \in V.$$

When $V = F^n$, quadratic forms are just a homogenous quadratic polynomial with n variables, i.e.

$$q(X) = X^t A X, \quad A \in F^{n\times n}, X \in F^n.$$

Let $Q(V)$ denote all the quadratic forms on V, it's an F-vector space. By definition there's a surjective linear map $M^2(V) \to Q(V)$ by $\Phi(f)(\alpha) = f(\alpha,\alpha)$.

Proposition 1.1.5
Let $\operatorname{char} F \neq 2$,

- The map $\Phi : S^2(V) \to Q(V)$ is an isomorphism.
- Let $q \in Q(V)$, if $f \in S^2(V)$ and $\Phi(f) = q$, then

$$f(\alpha,\beta) = \frac{1}{4}(q(\alpha+\beta) - q(\alpha-\beta)).$$

Proof. The first one can be proved by $\ker(\Phi) = \Lambda^2(V)$ and $M^2(V) = S^2(V) \oplus \Lambda^2(V)$.
The second one is trivial by direct computation. □

From this we can define the matrix of a quadratic form q to be the matrix of the symmetrical bilinear form $\Phi^{-1}(q)$, thus $[q]_{\mathcal{B}}$ is always symmetrical.

Theorem 1.1.6

Let $f \in M^2(V)$,

- If $\operatorname{char} F \neq 2$, then $f \in S^2(V) \iff \exists \mathcal{B}$, s.t. $[f]_{\mathcal{B}}$ diagonal;
- $f \in \Lambda^2(V) \iff \exists \mathcal{B}$ s.t. $[f]_{\mathcal{B}}$ is block diagonal with each block being 0 or $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

To prove this theorem, it's sufficient to prove:

Lemma 1.1.7

Let $f \in S^2(V) \cup A^2(V)$, $W \subset V$ is a subspace, let

$$W^{\perp} = \{\beta \in V \mid f(\alpha, \beta) = 0, \forall \alpha \in W\}.$$

If $f\big|_W$ is non-degenerate, then $V = W \oplus W^{\perp}$. In this case, let $\mathcal{B}_1, \mathcal{B}_2$ be basis of W, $W^{\perp}$, and $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$, we have

$$[f]_{\mathcal{B}} = \operatorname{diag}([f|_W]_{\mathcal{B}_1}, [f|_{W^{\perp}}]_{\mathcal{B}_2}).$$

Proof. Since $f\big|_W$ non-degenerate, $W \cap W^{\perp} = 0$. Note that

$$W^{\perp} = \bigcap_{\alpha \in W} \ker(L_f(\alpha)) = L_f(W)^0.$$

Thus $\dim W^{\perp} = n - \dim L_f(W) \geq n - \dim W$. This implies that $V = W \oplus W^{\perp}$.

For the second part, since $f(\alpha, \beta) = 0 \implies f(\beta, \alpha) = 0$, thus the matrix must obey the conclusion. □

Now by induction it's trivial when $n = 1$,

- When $f \in S^2(V)$, WLOG $f \neq 0$, $\exists \alpha$ s.t. $f(\alpha, \alpha) \neq 0$. Let $W = \operatorname{span}\{\alpha\}$, by lemma and induction hypo we're done.
- When $f \in A^2(V)$, there exists α, β s.t. $f(\alpha, \beta) = 1$. Let $W = \operatorname{span}\{\alpha, \beta\}$, similarly by lemma and induction hypo, we're done.

Corollary 1.1.8

For any $q \in Q(V)$, there exists a basis of V s.t. $[q]_{\mathcal{B}}$ diagonal.

The non-degenerate alternating bilinear forms are called **symplectic forms**.

Corollary 1.1.9

If there exists symplectic form f on V, then $\dim V = 2m$ and

$$[f]_{\mathcal{B}} = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}$$

for some basis $\mathcal{B}$.

Theorem 1.1.10

Let F be an algebraically closed field, and $\operatorname{char} F \neq 2$. Let $f \in S^2(V)$, there exists a basis $\mathcal{B}$, s.t. $[f]_{\mathcal{B}}$ diagonal and the diagonal entries can only be 0 or 1.

Proof. Use the previous result and multiply some scalars (the root of $x^2 = c$). $\square$

When $F = \mathbb{R}$, using similar technique we can prove the diagonal entries can only be $0, 1$ or -1.