Linear Algebra II

Felix Chen

Contents

Since this theorem requires the field to be algebraically closed, if T is in a smaller field, we wonder whether D and N is in that field.

Let $A \in \mathbb{R}^{n\times n}$, and $A = D + N$ be its Jordan decomposition. We'll prove that $D, N \in \mathbb{R}^{n\times n}$. By taking conjugates,

$$A = D + N \implies A = \overline{D} + \overline{N}.$$

It's clear that $\overline{D} + \overline{N}$ is also a Jordan decomposition of A, so we must have $D = \overline{D}$, which means $D \in \mathbb{R}^{n\times n}$.

In fact when $\mathbb{R}$ is replaced by any perfect field F, this property still holds. To prove this we need to introduce the semisimple maps.

§0.1 Semisimple transformations

As we've already seen, the "diagonalizable" property depends on the base fields, thus next we'll generalize the concepts of "diagonalizable".

Definition 0.1.1. Let $T \in L(V)$,

- We say T is **simple**(or irreducible) if V has no nontrivial T-invariant subspaces.
- We say T is **semisimple**(or totally reducible) if each T-invariant subspace $W \subset V$ there exists T-invariant subspace Z, s.t. $V = W \oplus Z$.

Obviously simple maps are always semisimple.

Proposition 0.1.2
Let T be a simple linear operator, then $\forall \alpha \in V\backslash\{0\}$, α is a cyclic vector of T.

Lemma 0.1.3

Let $T \in L(V)$.

- If T is semisimple, $V' \subset V$ is T-invariant, then $T_{V'}$ is semisimple.
- If $V = \bigoplus_{i=1}^k V_i$ s.t. T_{V_i} semisimple, then T is semisimple as well.

Proof. Suppose $W \subset V'$ is a T-invariant subspace. Since T is semisimple, $\exists Z \subset V$ s.t. $V = W \oplus Z$, and Z is T-invariant.

Let $Z' = Z \cap V'$, we claim that $V' = Z' \oplus W$.

Clearly $W \cap Z' = \{0\}$ and $W + Z' \subset V'$. For all $v \in V'$, $\exists w \in W, z \in Z$ s.t. $v = w + z$, since $v, w \in V'$, $z = v - w \in V'$ as well, which means $z \in Z'$.

For the second part, (We can assmue $k = 2$, but here we won't use it).

Let $W \subset V$ be a T-invariant subspace. Since T_{V_i} is semisimple, $\exists Z_i \subset V_i$ s.t.

$$V_i = \left(\left(W + \sum_{j=1}^{i-1} V_j \right) \cap V_i \right) \oplus Z_i.$$

Let $Z = \bigoplus_{i=1}^k Z_i$, we claim that $Z \oplus W = V$. If $w \in W \cap Z$, then $w = z_1 + \cdots + z_k$,

$$z_k = w - z_1 - \cdots - z_{k-1} \in Z_k \cap (W + V_1 + \cdots + V_{k-1}) = \{0\}.$$

Thus $z_k = 0$, similarly $z_{k-1} = \cdots = z_1 = 0 = w$.

Note that $W + \sum_{i=1}^j V_i \subset W \oplus \sum_{i=1}^j Z_i$ for all $j = 1, \ldots, k$, so $V = W \oplus Z$. □

Corollary 0.1.4

Let $T \in L(V)$, T is semisimple $\iff$ there exists a T-invariant decomposition $V = \bigoplus_{i=1}^k V_i$ s.t. each T_{V_i} is simple.

Theorem 0.1.5

Let $T \in L(V)$.

- T simple $\iff$ f_T is a prime polynomial;
- T semisimple $\iff$ p_T has no multiple factors.

Proof. T simple $\implies$ T cyclic $\implies$ $f_T = p_T$, so we only need to prove p_T is a prime.

Otherwise $p_T = gh$,

$$0 = p_T(T) = g(T)h(T),$$

So either $g(T)$ or $h(T)$ is not inversible. Thus $\ker(g(T)) \neq \{0\} \implies \ker(g(T)) = V \implies g(T) = 0$, contradiction!

If T is not simple, $\exists W \subset V$, W is T-invariant nontrivial subspace, so $f_T = f_{T_W} \cdot f_{T_{V/W}}$ is not a prime.

T semisimple $\implies \exists V_i$, $V = \bigoplus_{i=1}^k V_i$, such that T_{V_i} is simple $\implies p_{T_{V_i}}$ is prime.

$$p_T = \operatorname{lcm}(p_{T_{V_1}}, \ldots, p_{T_{V_k}})$$

has no multiple factors.

Conversely if p_T has no multiple factors, consider the primary cyclic decomposition of T :

$$V = \bigoplus_i W_i, \quad f_{T_{W_i}} \text{ primary.}$$

Since p has no multiple factors, $f_{T_{W_i}} = p_{T_{W_i}}$ is prime polynomial.

Hence T_{W_i} simple $\implies$ T semisimple. $\square$

Corollary 0.1.6

When F is an algebraically closed field:

- T simple $\iff$ $\dim V = 1$.
- T semisimple $\iff$ T is diagonalizable.

This corollary means that "semisimple" is indeed the equivalent description of "diagonalizable" in the algebraic closure.

Note that whether p_T has multiple factors or not does not change under *perfect* field extensions. So "semisimple" is a more general property (it stays the same under more transformations).

Recall that:

Definition 0.1.7 (Perfect fields)**.** If for all prime polynomials $p \in F[x]$, p has no multiple roots in $\overline{F}$, we say F is a **perfect field**.

Finite fields, fields with charcter 0 and algebraically closed fields are always perfect fields.

We can check that when F is perfect, $f \in F[x]$ has no multiple factors iff f has no multiple factors in $\overline{F}[x]$.

Now we can generalize the Jordan decomposition:

Theorem 0.1.8 (Jordan decomposition)

Let $T \in L(V)$, $n = \dim V$, and F is perfect. There exists unique $S, N \in L(V)$ s.t. $T = S + N$, where S semisimple and N nilpotent, and $SN = NS$.

Moreover there exists $f, g \in F[x]$ s.t. $S = f(T), N = g(T)$.

To prove this generalized version, we need the following observation:

Proposition 0.1.9

Let F be a perfect field, $A \in F^{n\times n}$ is semisimple iff A is diagonalizable in $\overline{F}^{n\times n}$.

Proof. A semisimple $\iff$ p_A has no multiple factors in $F[x]$

$\iff$ p_A has no multiple roots in $\overline{F}[x]$

$\iff$ p_A is the product of different monic polynomials of degree 1

$\iff$ A is diagonalizable in $\overline{F}^{n\times n}$. $\square$

Proposition 0.1.10

Let F be a perfect field, $a \in \overline{F}$. Then $a \notin F \iff$ exists an automorphism σ s.t. $\sigma\big|_F = \mathrm{id}_F$, i.e. $\sigma \in \mathrm{Gal}(\overline{F}/F)$ but $\sigma(a) \neq a$.

Remark 0.1.11 — This proof is beyond the scope of this class, but the idea is similar to the conjugate operation on $\mathbb{C}/\mathbb{R}$.

Now we prove the Jordan decomposition:

Proof. Let $A = S + N$ is the Jordan decomposition on $\overline{F}^{n\times n}$. Then by applying σ on this equation,

$$A = \sigma(S) + \sigma(N)$$

holds for all $\sigma \in \mathrm{Gal}(\overline{F}/F)$. Since $\sigma(S)$ is also diagonalizable, $\sigma(N)$ is nilpotent, as σ is an automorphism. So by the uniqueness of Jordan decomposition, $\sigma(S) = S, \sigma(N) = N$.

This implies $S, N \in F^{n\times n}$. □

§0.2 Bonus section

Starting from Galois groups mentioned above, let

$$\mathrm{Aut}(E/F) := \{\sigma \in \mathrm{Aut}(E) \mid \sigma|_F = \mathrm{id}_F\}$$

be the automorphism group of field extension E/F.

Example 0.2.1

Let $F = \mathbb{Q}$, $E = \mathbb{Q}(\sqrt{2})$, then $\sigma : a + b\sqrt{2} \mapsto a - b\sqrt{2}$ is in $\mathrm{Aut}(E/F)$.

If $E = \mathbb{Q}(\sqrt[3]{2})$, if $\sigma \in \mathrm{Aut}(E/F)$, then $\sigma(\sqrt[3]{2})$ is a root of $x^3 - 2 \implies \sigma = \mathrm{id}$. Thus E/F is not a *Galois extension*.

When E/F is a Galois extension, we write $\mathrm{Gal}(E/F) = \mathrm{Aut}(E/F)$.

In the history, this concept is used to solve polynomial equations.

Let $f \in \mathbb{Q}[x]$, let $x_1, \ldots, x_n$ be all roots of f. Consider $E = \mathbb{Q}(x_1, \ldots, x_n)$, and define $\mathrm{Gal}(f) = \mathrm{Gal}(E/\mathbb{Q})$. Back in the times of Galois, the concept of field haven't been developed yet, so what he did is to consider the bijections between the roots of f.

Galois discovered that f has radical solutions if and only if the group $\mathrm{Gal}(f)$ has a property, and he named it "solvable". Since all the subgroups of S_4 are solvable, thus if $\deg f \leq 4$, f always has radical solutions, but $A_5 < S_5$ is not solvable, so polynomials of degree greater than 4 may not have radical solutions.

One of the ultimate goal of modern algebra is to comprehend the group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

A tool developed for this goal is *group representation*. A representation of a group G is a homomorphism $\varphi : G \to \mathrm{GL}(V)$. Since $\mathrm{GL}(V)$ is something people knows very well, so when the elements of an abstract group G is viewed as linear maps, it's easier to discover more properties of G.

When $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, the representation is called a *Galois representation*. Even one dimensional Galois representations are very nontrivial.

Midterm exam QAQ

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

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 $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ 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
$$\langle \alpha, \beta \rangle = Y^* A X = \sum_{j,k=1}^{n} A_{kj} x_j \overline{y_k}$$
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)

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

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

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

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

Proposition 1.1.11 (Triangle inequality)

$$\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|.$$

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 $c_j = 0$, $\forall j$. $\square$

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

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. $\square$

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 $S = \{\alpha_1, \ldots, \alpha_m\} \subset V \backslash \{0\}$ 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. $\square$

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. $\square$