Math 672 Lecture 28

<u>Theorem</u> (the spectral theorem for real self-adjoint operators): Let V be a real inner product space and let $T \in \mathcal{L}(V)$ be self-adjoint. Then there exists an orthonormal basis $\mathcal{B}$ of V such that $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is diagonal.

<u>Proof:</u> We argue by induction on $\dim(V)$. If $\dim(V)=1$, the result is obviously true, so suppose it holds in every real inner product space of dimension $n-1$, where $n \geq 2$. Let V be a real inner product space of dimension n and let $T \in \mathcal{L}(V)$ be self-adjoint. By the preceding lemma, T has an eigenpair λ_1, v_1 ($\lambda_1 \in \mathbb{R}$). Assume, wlog, that $\|v_1\|=1$, and define $U = \text{span}(v_1)$. Then $V = U \oplus U^\perp$.
We know that $U^\perp$ is invariant under $T^* = T$. It is straightforward to show that $S = T|_{U^\perp}$ is a self-adjoint element of $\mathcal{L}(U^\perp)$. Hence, by the induction hypothesis, there exists an orthonormal basis $\{v_2, v_3, \ldots, v_n\}$ of $U^\perp$ and scalars $\lambda_2, \lambda_3, \ldots, \lambda_n \in \mathbb{R}$ such that

$$S(v_j) = \lambda_j v_j \text{ for } j = 2, 3, \ldots, n.$$

But then $T(v_j) = S(v_j) = \lambda_j v_j$ for $j = 2, 3, \ldots, n$ and hence $\{v_1, v_2, \ldots, v_n\}$

is an orthonormal basis for V with

$$T(v_j) = \lambda_j v_j, \quad j = 1, 2, \ldots, n.$$

This completes the proof by induction. //

Recall that the goal of the course is understand the "structure" of a linear operator on a finite-dimensional vector space. In practice this means answering the following question: How can we choose a basis $\mathcal{B}$ for V so that the matrix for $T \in \mathcal{L}(V)$ (w.r.t. $\mathcal{B}$) is as simple as possible? We have several results

- If V is complex, we can choose $\mathcal{B}$ so that $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is upper triangular. Moreover, it is possible to choose $\mathcal{B}$ to be orthonormal. (Schur's theorem)

- If V is complex and T is normal, then there exists an orthonormal matrix $\mathcal{B}$ for V such that $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is diagonal. We can then write

$$T(v) = \sum_{j=1}^{n} \lambda_j \langle v, v_j \rangle v_j \quad \forall\, v \in V \quad (\mathcal{B} = \{v_1, \ldots, v_n\})$$

  If T is not only normal but self-adjoint, then every λ_j is real (even though V is complex).

- If V is real and $T \in \mathcal{L}(V)$ is self-adjoint, we can choose an orthonormal basis $\mathcal{B}$ for V such that $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is diagonal.

  Again, we have

$$T(v) = \sum_{j=1}^{n} \lambda_j \langle v, v_j \rangle v_j,$$

  where now the λ_j's and v_j's are real.

- Otherwise, we only know that if $\dim(V) = n$ and there exist n linearly independent eigenvectors $v_1, v_2, \ldots, v_n \in V$ of T, then $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is diagonal ($\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$), and

$$T\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \lambda_j \alpha_j v_j.$$

  This is the case, in particular, if T has n distinct eigenvalues.

  This is true for any field F.

We now investigate the situation for a general linear operator $T \in \mathcal{L}(V)$, where V is a finite-dimensional vector space over F (usually $\mathbb{R}$ or $\mathbb{C}$; we get the most complete results when $F = \mathbb{C}$).

To look ahead: Let V be a complex vector space, let $T \in \mathcal{L}(V)$, and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. It is not always true that

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_k, T).$$

However, recalling that $E(\lambda_j, T) = \mathcal{N}(T - \lambda_j I)$, it is always true that

$$V = \mathcal{N}((T-\lambda_1 I)^{m_1}) \oplus \cdots \oplus \mathcal{N}((T-\lambda_k)^{m_k})$$

for some integers $m_1, \ldots, m_k$, and that

$$\mathcal{N}((T-\lambda_j I)^{m_j})$$

is invariant under T for each j. This is the basis of the Jordan canonical form of T.

We begin with some technical results. We express the first lemma in terms of a generic operator T, but we will eventually apply it to $T - \lambda_j I$.

Lemma: Let V be a finite-dimensional vector space over a field F and let $T \in \mathcal{L}(V)$. Then

1. $\{0\} = \mathcal{N}(T^0) \subseteq \mathcal{N}(T) \subseteq \mathcal{N}(T^2) \subseteq \cdots$
2. There exists m satisfying $0 \le m \le n = \dim(V)$ such that

$$\mathcal{N}(T^j) \subsetneq \mathcal{N}(T^{j+1}) \quad \forall\, 0 \le j < m \quad (\text{if } m > 0)$$

(where $\subsetneq$ means "is a proper subset of") and

$$\mathcal{N}(T^j) = \mathcal{N}(T^{j+1}) \quad \forall\, j \ge m.$$

<u>Proof:</u> By definition, $T^0 = I$, so $\mathcal{N}(T^0) = \{0\}$. If $v \in \mathcal{N}(T^j)$, then $T^j(v) = 0$ and

$$T^{j+1}(v) = T(T^j(v)) = T(0) = 0$$

$$\Rightarrow v \in \mathcal{N}(T^{j+1}).$$

Thus

$$\mathcal{N}(T^j) \subseteq \mathcal{N}(T^{j+1}) \quad \forall j \geq 0.$$

Now,

$$\mathcal{N}(T^j) \subseteq V \;\; \forall j$$

$$\Rightarrow \dim(\mathcal{N}(T^j)) \leq \dim(V) \;\; \forall j$$

and

$$\mathcal{N}(T^j) \subsetneq \mathcal{N}(T^{j+1}) \Rightarrow \dim(\mathcal{N}(T^{j+1})) \geq \dim(\mathcal{N}(T^j)) + 1.$$

It follows that

$$\mathcal{N}(T^j) \subsetneq \mathcal{N}(T^{j+1}) \;\; \forall j \geq 0$$

is impossible (otherwise, $\dim(\mathcal{N}(T^j)) > n = \dim(V)$ for j sufficiently large). Let $m \geq 0$ be the smallest integer such that

$$\mathcal{N}(T^{m+1}) = \mathcal{N}(T^m).$$

Note that $m \leq n$ (otherwise, $\dim(\mathcal{N}(T^{n+1})) > n$).

It remains only to show that
$$\mathcal{N}(T^{j+1}) = \mathcal{N}(T^j) \quad \forall j > m.$$
So let $j > m$ and let $v \in \mathcal{N}(T^{j+1})$. Then

$$\begin{aligned} T^{j+1}(v) = 0 &\Rightarrow T^{m+1}(T^{j-m}(v)) = 0 \\ &\Rightarrow T^{j-m}(v) \in \mathcal{N}(T^{m+1}) \\ &\Rightarrow T^{j-m}(v) \in \mathcal{N}(T^m) \quad (\text{since } \mathcal{N}(T^m) = \mathcal{N}(T^{m+1})) \\ &\Rightarrow T^m(T^{j-m}(v)) = 0 \\ &\Rightarrow T^j(v) = 0 \\ &\Rightarrow v \in \mathcal{N}(T^j). \end{aligned}$$

Thus $\mathcal{N}(T^{j+1}) \subseteq \mathcal{N}(T^j)$. Since we already know that $\mathcal{N}(T^j) \subseteq \mathcal{N}(T^{j+1})$, we see that $\mathcal{N}(T^{j+1}) = \mathcal{N}(T^j)$, as desired. //

<u>Lemma</u>: Let V be a finite-dimensional vector space over a field F and let $T \in \mathcal{L}(V)$. Suppose $\mathcal{N}(T^m) = \mathcal{N}(T^{m+1})$. Then

$$V = \mathcal{N}(T^m) \oplus \mathcal{R}(T^m).$$

<u>Proof</u>: By the fundamental theorem of linear algebra,

$$\dim(V) = \dim(\mathcal{N}(T^m)) + \dim(\mathcal{R}(T^m)).$$

Thus it suffices to prove that

$$\mathcal{N}(T^m) \cap \mathcal{R}(T^m) = \{0\}$$

Suppose $v \in \mathcal{N}(T^m) \cap \mathcal{R}(T^m)$. Then there exists $u \in V$ such that $v = T^m(u)$ and hence

$$\begin{aligned} v \in \mathcal{N}(T^m) &\Rightarrow T^m(T^m(u)) = 0 \\ &\Rightarrow T^{2m}(u) = 0 \\ &\Rightarrow T^m(u) = 0 \quad (\text{since } \mathcal{N}(T^{2m}) = \mathcal{N}(T^m)) \\ &\Rightarrow v = 0. \end{aligned}$$

This completes the proof. //

Definition: Let V be a vector space over a field F, let $T \in \mathcal{L}(V)$, and let $\lambda \in F$ be an eigenvalue of T. We say that $v \in V$ is a generalized eigenvector of T corresponding to λ iff $v \neq 0$ and there exists $j \geq 1$ such that

$$(T - \lambda I)^j (v) = 0.$$

We define the generalized eigenspace $G(\lambda, T)$ of T corresponding to λ to be the set of all generalized eigenvectors of T corresponding to λ, together with the zero vector. Thus $G(\lambda, T) = \mathcal{N}((T - \lambda I)^m)$, where

$$\mathcal{N}((T-\lambda I)^{m+1}) = \mathcal{N}((T-\lambda I)^m).$$

Note that every eigenvector of T is a generalized eigenvector of T, and $E(\lambda,T) \subseteq G(\lambda,T)$ for every eigenvalue λ.

In some cases, $E(\lambda,T) = G(\lambda,T)$; in fact, as we will see, if T is diagonalizable, then $E(\lambda,T) = G(\lambda,T)$ for every eigenvalue λ of T.

Theorem: Let V be a finite-dimensional vector space over a field F, let $T \in \mathcal{L}(V)$, let λ be an eigenvalue of T, and let $m \geq 0$ satisfy $\mathcal{N}((T-\lambda I)^{m+1}) = \mathcal{N}((T-\lambda I)^m)$. Then

$$V = \mathcal{N}((T-\lambda I)^m) \oplus \mathcal{R}((T-\lambda I)^m)$$
$$= G(\lambda,T) \oplus \mathcal{R}((T-\lambda I)^m)$$

and both $G(\lambda,T)$, $\mathcal{R}((T-\lambda I)^m)$ are invariant under T.

Proof: We already know that

$$V = \mathcal{N}((T-\lambda I)^m) \oplus \mathcal{R}((T-\lambda I)^m).$$

Suppose $v \in \mathcal{N}((T-\lambda I)^m)$. The

$$(T-\lambda I)^m(T(v)) = T((T-\lambda I)^m(v)) = T(0) = 0$$

(since polynomials in T commute) and hence $T(v) \in \mathcal{N}((T-\lambda I)^m)$.

This shows that $\mathcal{N}((T-\lambda I)^m)$ is invariant under T.

Now suppose $u \in \mathcal{R}((T-\lambda I)^m)$; then there exists $v \in V$ such that $u = (T-\lambda I)^m(v)$. But then

$$T(u) = T((T-\lambda I)^m(v)) = (T-\lambda I)^m(T(v)) \in \mathcal{R}((T-\lambda I)^m),$$

and hence $\mathcal{R}((T-\lambda I)^m)$ is also invariant under T. //