Generalized Eigenspace Decomposition

Lecture 9
Math 634
9/20/99

Eigenvalues don't have to be distinct for the analysis of the matrix exponential that was done last time to work. There just needs to be a basis of eigenvectors for $\mathbb{R}^n$ (or $\mathbb{C}^n$). Unfortunately, we don't always have such a basis. For this reason, we need to generalize the notion of an eigenvector.

First, some definitions:

Definition The *algebraic multiplicity* of an eigenvalue λ of an operator A is the multiplicity of λ as a zero of the characteristic polynomial $\det(A - xI)$.

Definition The *geometric multiplicity* of an eigenvalue λ of an operator A is the dimension of the corresponding eigenspace, *i.e.*, the dimension of the space of all the eigenvectors of A corresponding to λ.

It is not hard to show (*e.g.*, through a change-of-basis argument) that the geometric multiplicity of an eigenvalue is always less than or equal to its algebraic multiplicity.

Definition A *generalized eigenvector* of A is a vector x such that $(A-\lambda I)^k x = 0$ for some scalar λ and some $k \in \mathbb{N}$.

Definition If λ is an eigenvalue of A, then the *generalized eigenspace of A belonging to λ* is the space of all generalized eigenvectors of A corresponding to λ.

Definition We say that a vector space $\mathcal{V}$ is the *direct sum* of subspaces $\mathcal{V}_1, \dots, \mathcal{V}_m$ of $\mathcal{V}$ and write

$$\mathcal{V} = \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_m$$

if for each $v \in \mathcal{V}$ there is a unique $(v_1, \dots, v_m) \in \mathcal{V}_1 \times \cdots \times \mathcal{V}_m$ such that $v = v_1 + \cdots + v_m$.

Theorem (Primary Decomposition Theorem) *Let B be an operator on $\mathcal{E}$, where $\mathcal{E}$ is a complex vector space, or else $\mathcal{E}$ is real and B has real eigenvalues. Then*

$\mathcal{E}$ is the direct sum of the generalized eigenspaces of B. The dimension of each generalized eigenspace is the algebraic multiplicity of the corresponding eigenvalue.

Before proving this theorem, we introduce some notation and state and prove two lemmas.

Given $T : \mathcal{V} \to \mathcal{V}$, let

$$N(T) = \left\{ x \in \mathcal{V} \;\middle|\; T^k x = 0 \text{ for some } k > 0 \right\},$$

and let

$$R(T) = \left\{ x \in \mathcal{V} \;\middle|\; T^k u = x \text{ has a solution } u \text{ for every } k > 0 \right\}.$$

Note that $N(T)$ is the union of the null spaces of the positive powers of T and $R(T)$ is the intersection of the ranges of the positive powers of T. This union and intersection are each nested, and that implies that there is is a number $m \in \mathbb{N}$ such that $R(T)$ is the range of T^m and $N(T)$ is the nullspace of T^m.

Lemma *If $T : \mathcal{V} \to \mathcal{V}$, then $\mathcal{V} = N(T) \oplus R(T)$.*

Proof. Pick m such that $R(T)$ is the range of T^m and $N(T)$ is the nullspace of T^m. Note that $T|_{R(T)} : R(T) \to R(T)$ is invertible. Given $x \in \mathcal{V}$, let $y = \left(T|_{R(T)} \right)^{-m} T^m x$ and $z = x - y$. Clearly, $x = y + z$, $y \in R(T)$, and $T^m z = T^m x - T^m y = 0$, so $z \in N(T)$. If $x = \tilde{y} + \tilde{z}$ for some other $\tilde{y} \in R(T)$ and $\tilde{z} \in N(T)$ then $T^m \tilde{y} = T^m x - T^m \tilde{z} = T^m x$, so $\tilde{y} = y$ and $\tilde{z} = z$. □

Lemma *If λ_j, λ_k are distinct eigenvalues of $T : \mathcal{V} \to \mathcal{V}$, then*

$$N(T - \lambda_j I) \subseteq R(T - \lambda_k I).$$

Proof. Note first that $(T - \lambda_k I) N(T - \lambda_j I) \subseteq N(T - \lambda_j I)$. We claim that, in fact, $(T - \lambda_k I) N(T - \lambda_j I) = N(T - \lambda_j I)$; *i.e.*, that

$$(T - \lambda_k I)|_{N(T - \lambda_j I)} : N(T - \lambda_j I) \to N(T - \lambda_j I)$$

is invertible. Suppose it isn't; then we can pick a nonzero $x \in N(T - \lambda_j I)$ such that $(T - \lambda_k I)x = 0$. But if $x \in N(T - \lambda_j I)$ then $(T - \lambda_j I)^{m_j} x = 0$ for some $m_j \geq 0$. Calculating,

$$\begin{aligned}
(T-\lambda_j I)x &= Tx - \lambda_j x = \lambda_k x - \lambda_j x = (\lambda_k - \lambda_j)x, \\
(T-\lambda_j I)^2 x &= T(\lambda_k - \lambda_j)x - \lambda_j(\lambda_k - \lambda_j)x = (\lambda_k - \lambda_j)^2 x, \\
&\vdots \\
(T-\lambda_j I)^{m_j} x &= \cdots = (\lambda_k - \lambda_j)^{m_j} x \neq 0,
\end{aligned}$$

contrary to assumption. Hence, the claim holds.

Note that this implies not only that

$$(T-\lambda_k I)N(T-\lambda_j I) = N(T-\lambda_j I)$$

but also that

$$(T-\lambda_k I)^m N(T-\lambda_j I) = N(T-\lambda_j I)$$

for every $m \in \mathbb{N}$. This means that

$$N(T-\lambda_j I) \subseteq R(T-\lambda_k I).$$

□

Proof of the Principal Decomposition Theorem. It is obviously true if the dimension of $\mathcal{E}$ is 0 or 1. We prove it for $\dim \mathcal{E} > 1$ by induction on $\dim \mathcal{E}$. Suppose it holds on all spaces of smaller dimension than $\mathcal{E}$. Let $\lambda_1, \lambda_2, \dots, \lambda_q$ be the eigenvalues of B with algebraic multiplicities $n_1, n_2, \dots, n_q$. By the first lemma,

$$\mathcal{E} = N(B-\lambda_q I) \oplus R(B-\lambda_q I).$$

Note that $\dim R(B-\lambda_q I) < \dim \mathcal{E}$, and $R(B-\lambda_q I)$ is (positively) invariant under B. Applying our assumption to $B|_{R(B-\lambda_q I)} : R(B-\lambda_q I) \to R(B-\lambda_q I)$, we get a decomposition of $R(B-\lambda_q I)$ into the generalized eigenspaces of $B|_{R(B-\lambda_q I)}$. By the second lemma, these are just

$$N(B-\lambda_1 I), N(B-\lambda_2 I), \dots, N(B-\lambda_{q-1} I),$$

so

$$\mathcal{E} = N(B-\lambda_1 I) \oplus N(B-\lambda_2 I) \oplus \cdots \oplus N(B-\lambda_{q-1} I) \oplus N(B-\lambda_q I).$$

Now, by the second lemma, we know that $B|_{N(B-\lambda_k I)}$ has λ_k as its only eigenvalue, so $\dim N(B - \lambda_k I) \leq n_k$. Since

$$\sum_{k=1}^{q} n_k = \dim E = \sum_{k=1}^{q} \dim N(B - \lambda_k I),$$

we actually have $\dim N(B - \lambda_k I) = n_k$. $\square$