Math 108b: Notes 2/14/11

We have learned that not every linear operator T on a finite dimensional vector space is diagonalizable. In other words, we can't always find a basis of eigenvectors that makes $[T]_\beta$ a diagonal matrix. In fact, we know it's only possible when the characteristic polynomial splits and the dimensions of the eigenspaces are large enough!

Still, we would like to be able to take any linear operator T on a finite dimensional space and find a "good" basis – i.e., a basis in which $[T]_\beta$ is "almost" diagonal. (We will still need the characteristic polynomial to split to have any hope of being able to do this.) What should "almost" diagonal mean? Recall our simple example of a matrix that is not diagonalizable to motivate the following definition:

DEFINITION A matrix $J \in M_{n\times n}(F)$ is in **Jordan canonical form** if

$$J = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix},$$

where each A_i is a *Jordan block matrix* – that is, each A_i is a square matrix with a constant λ_i on the diagonal and 1's on the off-diagonal:

$$A_i = \begin{pmatrix} \lambda_i & 1 & 0 & ... & 0 & 0 \\ 0 & \lambda_i & 1 & ... & 0 & 0 \\ \vdots & & & & \vdots & \\ 0 & 0 & 0 & ... & \lambda_i & 1 \\ 0 & 0 & 0 & ... & 0 & \lambda_i \end{pmatrix}$$

EXAMPLES

1. A diagonal matrix is in Jordan canonical form – each block is just the 1×1 matrix (a_{ii}).

2. Locate each block matrix in the following matrix A, which is in Jordan canonical form:

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

Notice that we can easily compute the characteristic polynomial of a matrix in Jordan form. For example, for the above matrix, $f(t) = \det(A - tI) = (2-t)^4(3-t)^3(5-t)$. The diagonal elements are therefore the eigenvalues! We will see that we can actually tell from looking at the Jordan canonical form what the dimension of each eigenspace is! For instance, for A, $\dim E_2 = 3$, $\dim E_3 = 1$ and $\dim E_5 = 1$. For the eigenvalue 2, name three linearly independent eigenvectors: ________________.

Also, ______ is an eigenvector with eigenvalue 3 and ______ is an eigenvector with eigenvalue 5. Therefore, we've found five linearly independent eigenvectors – unfortunately, it's fairly easy to show we can't find any more. Therefore, we don't have a basis of eigenvectors; this means A cannot be diagonalized – the best we can do is leave it in the Jordan canonical form.

Jordan block matrices have a special property: If A has λ on the diagonal and 1's on the off-diagonal, when we look at $A - \lambda I$, we have a matrix with zeros everywhere, except on the off-diagonal! What happens when we raise this matrix to a power? E.g.,

When A is a Jordan block matrix, $(A - \lambda I)^p =$ ___ (the ________________) for some $p \in \mathbb{N}$.

DEFINITIONS: Let T be a linear operator on a vector space V

A **generalized eigenvector** is a vector $v \in V$ such that _______ and ________________ for some $p \in \mathbb{N}$.

The **generalized eigenspace** corresponding to an eigenvalue λ is

$$K_\lambda = \{x \in V : \underline{\hspace{8cm}}\}$$

THEOREM Properties of K_λ.
Let $T \in \mathcal{L}(V)$, and let λ be an eigenvalue of T with multiplicity m.

(a) K_λ is a subspace of V (and, clearly, $K_\lambda \supseteq E_\lambda$.)
(b) K_λ is a T-invariant subspace of V.
(c) Define the operator $S = T - \mu I$. If $\mu \neq \lambda$, then S_{K_λ} (that is, the restriction of S to the T-invariant space K_λ) is one-to-one.
(d) $\dim(K_\lambda) \leq m$
(e) $K_\lambda = N((T - \lambda I)^m)$

Proof (a) Clearly $0 \in K_\lambda$, and K_λ is closed under scalar multiplication. We need only show that K_λ is closed under vector addition:

(b) Let $v \in K_\lambda$. This means that ________________________________. Therefore,

(c) To prove this, show that the null space of S_{K_λ} contains only the 0 vector. For a contradiction, assume that

Part (d) follows since the characteristic polynomial of T_{K_λ} is $g(t) = (\lambda - t)^d$ (where $d = \dim K_\lambda$) divides the characteristic polynomial of T. By definition of multiplicity, $d \leq m$. Part (e) follows from the Cayley-Hamilton theorem: $g(T_{K_\lambda}) = T_o$. Therefore, for all $x \in K_\lambda$, $(\lambda I - T_{K_\lambda})^d x = (\lambda I - T)^d x = 0$. Hence, $K_\lambda \subseteq N((T - \lambda I)^d) \subseteq N((T - \lambda I)^m)$ since $d \leq m$. (From the definition of K_λ, it is clear that $N((T - \lambda I)^m) \subseteq K_\lambda$.)

THEOREM If T is a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits, then there exists a basis β such that $[T]_\beta$ is in Jordan canonical form.

The basic idea is that if the characteristic polynomial splits, and so has k distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$ with corresponding mulitplicities $m_1, m_2, ..., m_k$ that sum up to $n = dim(V)$, then the generalized eigenspaces K_{λ_j} will each have the full dimension $\dim K_{\lambda_j} = m_j$. Then, we can choose a basis β_j of *generalized* eigenvectors for each generalized eigenspace, and the union of these will be a basis of *generalized* eigenvectors for all of V.

Of course, $[T]_\beta$ will only be diagonal if each $K_\lambda = E_\lambda$, so all the generalized eigenvectors are true eigenvectors. This is the only way we will have a basis of all eigenvectors!

It is fairly easy to prove (see Theorems 7.3, 7.4 in your book) that there is a basis of generalized eigenvectors. The only problem is that we still have to choose each basis β_j of K_{λ_j} carefully in order to make sure that $[T]_\beta$ will be in Jordan canonical form!

DEFINITIONS Let T be a linear operator on a vector space V and let λ be an eigenvalue of T. Let x be a generalized eigenvector of T corresponding to the eigenvalue λ and let p be the smallest integer such that $(T - \lambda I)^p(x) = 0$. The set

$$\Big\{ \qquad\qquad\qquad\qquad\qquad\qquad \Big\}$$

is the **cycle of generalized eigenvectors** of T corresponding to λ with **initial vector** x. The **length** of the cycle is p.

Consider Example 2. For $\lambda = 3$, note that e_5, e_6, and e_7 are all examples of generalized eigenvectors. If we use e_5 as the initial vector, then the cycle is simply $\{e_5\}$ (since $(A-3I)e_5 = 0$). If we use e_6 as the initial vector, we find the cycle $\{e_5, e_6\}$ (since $(A-3I)e_6 = e_5$). Finally, if we use e_7, we find the cycle $\{e_5, e_6, e_7\}$. This is of course the cycle we wish to use as our basis for K_3! It has the right length, and the first element of it is a true eigenvector.

For $\lambda = 2$, we would have that $\{e_1, e_2\}$ is a cycle of generalized eigenvectors, as is $\{e_3\}$ and $\{e_4\}$. We have to union these three cycles together to find a good basis for K_2 – this is because there are three separate Jordan blocks for the eigenvalue 2.

THEOREM Let $T \in \mathcal{L}(V)$, and let λ be an eigenvalue of T.

(a) Every cycle of generalized eigenvectors of T is linearly independent.

(b) Assume V is finite dimensional. Then, K_λ has an ordered basis that is a union of disjioint cycles of generalized eigenvectors corresponding to λ. In this basis, $[T_{K_\lambda}]$ is in Jordan canonical form.

Proof: The proof of part (a) follows by induction on the length of the cycle, and the proof of part (b) follows by induction on the dimension of K_λ.

(a) Every cycle of length 1 is linearly independent (note that in this case, the initial vector must have been an eigenvector). Now assume that every cycle of length $n-1$ is linearly independent. Assume that x is a generalized eigenvector that generates a cycle of length n. (Note this means that $(T-\lambda I)^{n-1}(x) \neq 0$ and $(T-\lambda I)^n x = 0$.) We want to show that

$$\left\{ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \right\}$$

is linearly independent.

Letting $y =$ ______________, we see that the cycle generated by y is

$$\left\{ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \right\},$$

which has length $n-1$. Therefore, by assumption it is a linearly independent set. Now, assume that

$$(\star)$$

for some scalars $a_0, a_1, ...a_{n-1}$. We must show that these scalars are all 0. Apply ____________ to the above equation:

Since ______________________________________ is linearly independent, we must have that $a_0 = ... = a_{n-2} = 0$. Then, $(\star)$ becomes $a_{n-1}(T-\lambda I)^{n-1}x = 0$, and since $(T-\lambda I)^{n-1}(x) \neq 0$, we have that $a_{n-1} = 0$. Hence, we have shown that any cycle of length n is linearly independent. $\square$