Jordan Normal Form

April 24, 2007

Definition: A Jordan block is a square matrix $B$ whose diagonal entries consist of a single scalar $\lambda$, whose superdiagonal entries are all 1, and all of whose other entries vanish. For example:

$$
\begin{pmatrix}
\lambda & 1 & 0 & 0 & \cdots & 0 \\
0 & \lambda & 1 & 0 & \cdots & 0 \\
0 & 0 & \lambda & 1 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & 0 & \lambda
\end{pmatrix}
$$

Theorem: Let $T$ be a linear operator on a finite dimensional vector space $V$. Suppose that the characteristic polynomial of $V$ splits. Then there exists a basis for $T$ such that $[T]_\beta$ is a direct sum of Jordan blocks.

The first step in the proof of this theorem is to use the direct sum decomposition of $V$ into generalized eigenspaces $K_\lambda$. Then it suffices to prove the theorem for the restriction of $T$ to each $K_\lambda$. On $K_\lambda$, let $S_\lambda := T - \lambda I$. If we can find a basis $\beta$ of $K_\lambda$ with respect to which $S_\lambda$ is a sum of Jordan blocks, then the same will be true for $T$. On $K_\lambda$, there exists an $r$ such that $S_\lambda^r = 0$. Thus it suffices to consider the special case of operators with this property.

Let $V$ be a finite dimensional vector space over a field $F$. A linear operator $N: V \to V$ is said to be nilpotent if $N^r = 0$ for some positive integer $r$. Let $N$ be a nilpotent operator on a finite dimensional vector space $V$. For each $i$, let $R^i$ be the image of $N^i$. Each $R^i$ is a linear subspace of $V$ and is $N$-invariant, and $0 = R^r \subseteq R^{r-1} \subseteq \ldots \subseteq R^1 \subseteq V$. Since $N$ is nilpotent it is not injective (unless $V = 0$). Thus the kernel $K$ of $N$ is not zero and $\dim R^1 = \dim V - \dim K < \dim V$.

Let $(v_1, v_2, \cdots, v_s)$ be a basis for $V$. Then $[N]_\beta$ is a Jordan block if and only if $N(v_1) = 0$, $N(v_2) = v_1$, and $N(v_i) = v_{i-1}$ for all $i > 1$. This motivates the
following definition.

**Definition:** An \( N \)-cycle is a sequence \( (v_1, v_2, \cdots, v_s) \) of nonzero vectors such that \( N(v_i) = v_{i-1} \) for all \( i > 1 \) and \( N(v_1) = 0 \).

If \( (v_1, \cdots, v_s) \) is an \( N \)-cycle, then \( v_1 = N^{s-1}(v_s) \), so \( v_1 \in R^{s-1} \). Conversely, if \( v \in R^{s-1} \), say \( v = R^{s-1}(x) \), then \( (R^{s-1}(x), R^{s-2}(x), \cdots, x) \) is an \( N \)-cycle whose initial vector is \( v \). If \( v \) belongs to \( R^s \) but not to \( R^r \), then \( s \) is the length of the longest \( N \)-cycle starting with \( v \).

**Definition:** An \( N \)-cycle \( (v_1, \cdots, v_s) \) is maximal if \( v_1 \not\in R^s \).

It is clear that every nonzero element of the kernel \( K \) of \( N \) is contained in some maximal \( N \)-cycle.

**Lemma:** Let \( (\gamma_1, \gamma_2, \cdots, \gamma_p) \) be a sequence of \( N \)-cycles. Then if the corresponding sequence of initial vectors is linearly independent, so is the concatenated sequence \( \gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_p \).

**Proof:** Say \( \gamma_i = (v_{i,1}, v_{i,2}, \cdots, v_{i,n_i}) \). Our assumption is that the sequence \( (v_{1,1}, v_{2,1}, \cdots, v_{p,1}) \) is linearly independent, and we want to prove that the entire (multi-indexed) sequence \( (v_{i,j}) \) is linearly independent. We prove this by induction on the maximum of the \( n_i \)'s. If all the \( n_i \)'s are 1, there is nothing to prove, since we assumed that the sequence of initial vectors is linearly independent. For the induction step, for each \( i \) let \( \gamma_i' \) be the (possibly empty) Jordan cycle obtained by omitting the last term. The induction assumption says that the union of these is linearly independent. Suppose that \( \sum a_{i,j}v_{i,j} = 0 \). Applying \( N \), we deduce that \( \sum a_{i,j}Nv_{i,j} = 0 \), i.e., that \( \sum a_{i,j}v_{i-1,j} = 0 \), where here for each \( j \), \( i \) ranges between 2 and \( n_i \). This is the sum over the corresponding truncated cycles \( \gamma_i' \). The induction assumption says that \( \cup \gamma_i' \) is linearly independent, so \( a_{i,j} = 0 \) for \( i \geq 2 \). Thus the original sum reduces to a linear combination of the initial vectors, which we assumed to be linearly independent. Hence each \( a_{1,j} = 0 \) as well.

Recall that we have linear subspaces \( 0 \subseteq R^r \subseteq R^{r-1} \subseteq \cdots \subseteq V \). Consider the corresponding sequence of subspaces of \( K \).

\[
0 = R^r \cap K \subseteq R^{r-1} \cap K \subseteq \cdots \subseteq R^1 \cap K \subseteq K.
\]

We shall say that a basis \( \alpha \) of \( K \) is adapted to \( N \) if for each \( i \), \( \alpha \cap R^i \) is a basis of \( R^i \cap K \). It is clear that such bases always exist: start with a basis for \( R^{r-1} \), extend it to a basis for \( R^{r-2} \), and continue.

**Definition:** A sequence of maximal \( N \)-cycles \( (\gamma_1, \cdots, \gamma_q) \) is full if the corresponding sequence of initial vectors \( (v_1, \cdots, v_q) \) is a basis of \( K \) which is adapted to \( N \).
It is clear that full sequences of $N$-cycles exist: start with a basis for $K$ which is adapted to $N$, and for each vector $v$ in the basis, find a a maximal cycle starting with $v$.

**Theorem:** Every full sequence of maximal $N$-cycles forms a basis for $V$.

**Proof:** Let $(\gamma_1, \gamma_2, \cdots, \gamma_p)$ be a full sequence of maximal $N$-cycles. By assumption, the corresponding sequence of initial vectors is linearly independent, and hence by the lemma, the concatenation of $\gamma_i$’s is linearly independent. It suffices to show that it also spans $V$. We do this by induction on the smallest $r$ such that $N^r = 0$. If $r = 1$, then $V = K$ and there is nothing to prove, since we assumed that the initial vectors span $K$. Let $V' := \text{Im}(N)$ and for each $i$, let $\gamma'_i$ be $\gamma_i$ with the last element omitted. In fact, $\gamma'_i = N(\gamma_i)$, with zero omitted. Let $N'$ be the restriction of $N$ to $V'$. Each $\gamma'_i$ is contained in $V'$ and is a maximal Jordan cycle for $N'$. Furthermore, $\gamma'_i$ is empty only if $\gamma_i$ has length one, which is true only if its initial (and only) vector does not belong to $V'$. Thus the sequence of initial vectors of $\gamma'_i$ contains all the initial vectors of the original sequence which belong to $V'$. Let $p'$ be the number of nonempty $\gamma'_i$’s. It follows that the sequence $(\gamma'_1, \cdots, \gamma'_p)$ is maximal and full for $N'$. By the induction assumption, it spans $V'$. Now let $W$ be the span of the all the $\gamma_i$’s. Note that by construction, $W$ contains all of $K$. Furthermore, the image of $W$ under $N$ contains all the $\gamma_i$’s and hence all of $V' = \text{Im}(N)$. But then $\dim W = \dim K + \dim \text{Im}(N) = \dim V$, and hence $W = V$.

**Remark:** For each $i$, let $d_i$ denote the dimension of $R^i$ and let $h_i := d_{i-1} - d_i$. If $\alpha$ is any basis for $K$ adapted to $N$, then $d_i$ is the number of elements of $\alpha$ which lie in $R^i$ and so $h_i$ is the number of elements of $\alpha$ which lie in $R^{i-1}$ but not in $R^i$. Corresponding to each such element there will be a maximal $N$-cycle of length $i$. Thus if $\beta$ is the basis obtained as above, the corresponding matrix $[N]_\beta$ will have exactly $h_i$ Jordan blocks of length $i$.

Let $V$ and $V'$ be two finite dimensional vector spaces over $F$, and let $T$ be an operator on $V$ and $T'$ an operator on $V'$. Then $T$ and $T'$ are sometimes said to be similar if there exists an isomorphism $S:V \rightarrow V'$ such that $T' \circ S = S \circ T$, i.e., $T' = S \circ T \circ S^{-1}$.

**Theorem:** Suppose that $f_T(x)$ and $f_{T'}(x)$ split. Choose bases $\beta$ for $V$ and $\beta'$ for $V'$ such that $A := [T]_\beta$ and $A' := [T']_{\beta'}$ are direct sums of Jordan blocks. Then $T$ and $T'$ are similar if and only if for each $\lambda \in F$ and each integer $s$, the number of Jordan blocks of $A$ with eigenvalue $\lambda$ and length $s$ is the same as the corresponding number for $A'$. 
