Notes on the Jordan Canonical Form

1 The Preliminary Block-Diagonal Form

In the first section, we will assume that a similarity transformation has been made so that a given matrix has been put into upper-triangular form. We then show how to further reduce the matrix, using a similarity transformation, to a block-diagonal form where the blocks on the diagonal are each upper-triangular.

**Theorem 1.1** Suppose that an upper-triangular matrix $U$ has the following form:

$$
U = \begin{pmatrix}
U_{11} & U_{12} & \cdots & U_{12} \\
O & U_{22} & \cdots & U_{2s} \\
\vdots & \vdots & \ddots & \vdots \\
O & O & \cdots & U_{ss}
\end{pmatrix}
$$

where each $U_{ii}$ is an upper-triangular matrix, all of whose diagonal elements are equal to $\lambda_i$. Also, suppose that $\lambda_1, \ldots, \lambda_s$ are all distinct. Then there exists a non-singular matrix $R$ such that

$$
V = R^{-1} U R = \begin{pmatrix}
V_1 & O & \cdots & O \\
O & V_2 & \cdots & O \\
\vdots & \vdots & \ddots & \vdots \\
O & O & \cdots & V_s
\end{pmatrix}
$$

where $V_i$ is an upper-triangular matrix of the same size as $U_{ii}$ and all of whose diagonal elements are equal to $\lambda_i$. 
We will refer to this latter form as the block-diagonal upper-triangular form.

**Proof:** Let \( Q = I + K \) where \( K \) is a matrix all of whose elements are zero except the \((p,q)\) element which is equal to \( k \), where \( p < q \). Note that \( Q \) is an elementary matrix derived from the identity matrix by adding \( k \) times the \( q^{th} \) row to the \( p^{th} \) row. Its inverse is then the elementary matrix \( I - Q \).

Recall that pre-multiplication of any matrix by an elementary matrix automatically carries out an elementary operation on the rows of the matrix, while post-multiplication carries out the elementary operation on the columns. Hence, pre-multiplication by \( Q^{-1} \) replaces the \( p^{th} \) row of \( U \) with row \( p - k \times \) row \( q \). Since the element in the \((p,q)\)-position lies in the \( q^{th} \) column, the entry \( u_{pq} \) is replace by \( u_{pq} - k u_{qq} \) and modifies elements in the \( p^{th} \) row only to the right of the entry \( u_{pq} \).

Let us pause in the proof to give an example:

**Example 1.2** Consider the matrix

\[
A = \begin{pmatrix}
2 & 4 & 1 & 6 \\
0 & 2 & 5 & 1 \\
0 & 0 & 4 & 1 \\
0 & 0 & 0 & 4 \\
\end{pmatrix}
\]

and take

\[
I - K = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & -k & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

Then, as is easily checked
\[(I - K)A = \begin{pmatrix} 2 & 4 & 1 & 6 \\ 0 & 2 & 5 - 4k & 1 - k \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}.\]

So the \((2, 3)\) element is \(a_{pq} - k (a_{qq}) = a_{23} - k (a_{33}) = 5 - k (4).\)

Now post-multiplication by \(I + K\) will add \(k\times\) col \(p\) to col \(q\). Thus the elements in the \(q^{th}\) column are replaced by col \(q + k\) col \(p\). This means that \(u_{pq}\) is replaced by \(u_{pq} + k u_{pp}\) and only entries above the \(p^{th}\) row are modified.

Returning to the example:

**Example 1.3** With \(A\) as above we compute the product

\[A(I + K) = \begin{pmatrix} 2 & 3 & 4k + 1 & 6 \\ 0 & 2 & 5 + 2k & 1 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \]

Note that now \(a_{pq} - k (a_{pp}) = a_{23} + k (a_{22}) = 5 + k (2).\)

Combining these results, we find that the similarity transform \(Q^{-1}UQ\) transforms \(u_{pq}\) \((p < q)\) into

\[u_{pq} - k (u_{qq} - u_{pp}),\]

and otherwise modifies only elements in \(U\) in the \(p^{th}\) row to the right of \(u_{pq}\) and in the \(q^{th}\) column above \(u_{pq}\). If \(u_{pp} \neq u_{qq}\) then we can choose \(k = u_{pq}/(u_{qq} - u_{pp})\) so that the similarity transformation replaces \(u_{pq}\) with 0.
Referring to the original form of the matrix $U$ since each of the $U_{ii}$ have the same diagonal entries, this allows us by means of a similarity transformation, to remove all the entries in the matrices $U_{pi}$ above the matrix $U_{ii}$, one entry at a time. The resulting form is that which is required.

**Example 1.4** In terms of the concrete example, the matrix product

$$(I - K) A (I + K) = \begin{pmatrix}
2 & 4 & 4k + 1 & 6 \\
0 & 2 & 5 - 4k + 2k & 1 - k \\
0 & 0 & 4 & 1 \\
0 & 0 & 0 & 4
\end{pmatrix}$$

we compute $a_{pq} - k(a_{qq} - a_{pp}) = a_{33} - k(a_{33} - a_{22}) = 5 - k(4 - 2) = 5 - 2k$ So taking $k = 5/2$, the new entry in the $(pq)$ position is zero. Indeed, substituting $k = 5/2$ in the preceding matrix yields

$$(I - K) A (I + K) = \begin{pmatrix}
2 & 4 & 11 & 6 \\
0 & 2 & 0 & -3/2 \\
0 & 0 & 4 & 1 \\
0 & 0 & 0 & 4
\end{pmatrix}$$

Likewise, the entry in the $(1,4)$ position in the matrix $A$ is 6 while the entries $u_{44} = 4$ and $u_{11} = 2$. Hence taking the value of $k = 6/2 = 3$ we find that $a_{14} - k(a_{44} - a_{11}) = 6 - 3(4 - 2) = 0$. Now the matrix looks like

$$\begin{pmatrix}
2 & 4 & 11 & 6 - 4k + 2k \\
0 & 2 & 0 & -3/2 \\
0 & 0 & 4 & 1 \\
0 & 0 & 0 & 4
\end{pmatrix}$$

So, after two iterations, the similar matrix is
\[ \tilde{A} = \begin{pmatrix} 2 & 4 & 11 & 0 \\ 0 & 2 & 0 & -3/2 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \]

It remains to remove the entries 11 in the (1,3) position and the $-\frac{3}{2}$ in the (2,4) position. Having done this, the matrix will take the form

\[ V = \begin{pmatrix} 2 & 4 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \]

which is in block-diagonal upper-triangular form, as described in the theorem.

Note that the (1,2) entry and the (3,4) elements cannot be reduced to zero since the entries in the (1,1) and (2,2) entries are equal as are the entries in the (3,3) and (4,4) positions as well.

**Exercise 1.5** Reduce the following upper-triangular matrix to block-diagonal form by means of a similarity transformation.

\[ A = \begin{pmatrix} -2 & 3 & 4 \\ 0 & 4 & 2 \\ 0 & 0 & 4 \end{pmatrix} \]

To summarize the work so far, we can start with a matrix $A$, reduce it to Schur upper-triangular form $T = P^{-1} A P$, and then reduce $T$ to block-diagonal, upper-triangular form $V = (PR)^{-1} A (PR)$ so that, with respect to the basis formed from the columns of $PR$, the linear transformation $A$
defined by $A(x) = Ax$ is represented by the simplified matrix $V$. Moreover, if we take an arbitrary vector $v = (y_1, y_2, \cdots, y_s)^T$ where $y_i$ is an $r$-vector if $V_i$ is $r \times r$. Then the matrix $A$ takes a vector of the form $(O, O, \cdots, O, y_i, O, \cdots, O)^T$ into a vector of the same form. In this way, the similarity transformation we have derived decomposes the linear transformation $A$ into $s$ different “mini”-transformations.

In fact, it is possible to simplify even further. Each matrix $V_i$ has the form

$$V = \begin{pmatrix}
\lambda & v_{12} & \cdots & v_{1q} \\
0 & \lambda & \cdots & v_{2q} \\
& & \ddots & \\
0 & 0 & \cdots & \lambda
\end{pmatrix}.$$ 

In fact, it is possible to find a similarity transformation that will put $V$ into a simpler block-diagonal, upper-triangular form where the blocks are called Jordan blocks. Thus we can find a non-singular matrix $S$ such that

$$S^{-1} VS = \begin{pmatrix}
J_1 & O & \cdots & O \\
O & J_2 & \cdots & O \\
& & \ddots & \\
O & O & \cdots & J_q
\end{pmatrix},$$

where the diagonal elements are all equal to $\lambda$ and where each $J_i$ is a Jordan block which is a square matrix whose elements are zero except for those on the principal diagonal, which are all equal, and those on the first superdiagonal which are all equal to one. Such a matrix then looks like
\[
J = \begin{pmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
0 & \lambda & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \lambda
\end{pmatrix}.
\]

This fact, coupled with the previous results, leads to the following statement.

**Theorem 1.6 (Jordan Canonical Form)** If \(A\) is a general square \(n \times n\) matrix then a non-singular matrix \(Q\) exists such that

\[
Q^{-1}AQ = \begin{pmatrix}
J_1 & 0 & \cdots & 0 \\
0 & J_2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & J_k
\end{pmatrix} = J
\]

where the \(J_i\) are \(n_i \times n_i\) Jordan blocks. The same eigenvalues may occur in different blocks, but the number of distinct blocks corresponding to a given eigenvalue is equal to the number of eigenvectors corresponding to that eigenvalue and forming an independent set. The number \(k\) and the set of numbers \(n_1, n_2, \ldots, n_k\) are uniquely determined by \(A\).

We will explore how to compute the Jordan canonical form in the next section.

**Exercise 1.7** Reduce the upper triangular matrix

\[
A = \begin{pmatrix}
1 & -2 & 3 & -4 \\
0 & 1 & -1 & -2 \\
0 & 0 & 1 & 4 \\
0 & 0 & 0 & -3
\end{pmatrix}
\]
2 Computing the Jordan Form

The relation $Q^{-1} A Q = J$ is equivalent to the relation $A Q = Q J$ and a careful look at the form of $J$ shows us that if we denote the columns of $Q$ by $q$ then the equation $A Q = Q J$ separates into a series of equations of the form

$$A q_i = \lambda_i q_i + \nu_i q_{i-1},$$

where $\nu_i$ may be either 0 or 1, depending on the superdiagonal entry of $J$.

Let us suppose, following our example, that

$$A = \begin{pmatrix} 2 & 4 & 1 & 6 \\ 0 & 2 & 5 & 1 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \quad \text{and its Jordan Form} \quad V = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

The eigenvalues of $A$ are $\lambda_1 = 2$ which has multiplicity 2 and $\lambda_2 = 4$ which also has multiplicity 2. Then the equation for the eigenvector or eigenvectors associated with $\lambda_1$ are given by the equation

$$\begin{pmatrix} 0 & 4 & 1 & 6 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
\[ 4v_2 + v_3 + 6v_4 = 0 \]
\[ 5v_3 + v_4 = 0 \]
\[ 2v_4 = 0 \]

hence \( v_2 = v_3 = v_4 = 0 \) and we take \( q_1 = (1, 0, 0, 0)^T \) as the required eigenvector. There is no other eigenvector. Note that

\[
A q_1 = \begin{pmatrix}
2 & 4 & 1 & 6 \\
0 & 2 & 5 & 1 \\
0 & 0 & 4 & 1 \\
0 & 0 & 0 & 4 \\
\end{pmatrix} \begin{pmatrix}
u_1 \\
u_2 \\
u_3 \\
u_4 \\
\end{pmatrix} = \begin{pmatrix}
1 \\
0 \\
0 \\
0 \\
\end{pmatrix} = 2 \begin{pmatrix}
1 \\
0 \\
0 \\
0 \\
\end{pmatrix} = 2q_1.
\]

as required.

Now we look at the equation involving \( q_2 \).

\[
A q_2 = 2q_2 + q_1
\]

\[
= \begin{pmatrix}
2 & 4 & 1 & 6 \\
0 & 2 & 5 & 1 \\
0 & 0 & 4 & 1 \\
0 & 0 & 0 & 4 \\
\end{pmatrix} \begin{pmatrix}
u_1 \\
u_2 \\
u_3 \\
u_4 \\
\end{pmatrix} = \begin{pmatrix}
10 \\
4 \\
0 \\
0 \\
\end{pmatrix} = 2 \begin{pmatrix}
4 \\
2 \\
0 \\
0 \\
\end{pmatrix} + 1 \begin{pmatrix}
1 \\
0 \\
0 \\
0 \\
\end{pmatrix}.
\]

This equation is then equivalent to the system

\[
4u_2 + u_3 + 6u_4 = 1 \\
5u_3 + u_4 = 0 \\
2u_4 = 0 \\
\]

and we may take \( u_1 \) arbitrarily. Choosing \( u_1 = 0 \) we then have \( u_2 = 1/4 \)
and \( u_3 = u_4 = 0 \). Then the vector \( q_2 = (0, 1/4, 0, 0)^\top \) is the generalized eigenvector associated with \( \lambda_1 = 2 \).

We now turn to the next eigenvalue \( \lambda_2 = 4 \). Then

\[
A - 4I = O \quad \text{becomes} \quad \begin{pmatrix} -2 & 4 & 1 & 6 \\ 0 & -2 & 5 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\]

It is easy to see that the corresponding eigenvector is (up to a constant multiple) \( q_3 = (11, 5, 2, 0)^\top \). We now look for a solution of

\[
A q_4 = 4 q_4 = q_3.
\]

Again, writing out the corresponding system leads to

\[
\begin{align*}
-2u_1 + 4u_2 + u_3 + 6u_4 &= 11 \\
-2u_2 + 5u_3 + u_4 &= 5 \\
u_4 &= 2.
\end{align*}
\]

Setting \( u_1 = 0 \) and solving for \( u_2 \) and \( u_3 \) yields the generalized eigenvector \( q_4 = (0, -4/11, 5/11, 2)^\top \).

Now we form the matrix \( Q = \col [q_1, q_2, q_3, q_4] \) and compute its inverse \( Q^{-1} \).

In this case

\[
Q = \begin{pmatrix} 1 & 0 & 11 & 0 \\ 0 & 1/4 & 5 & -4/11 \\ 0 & 0 & 2 & 5/11 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad Q^{-1} = \begin{pmatrix} 1 & 0 & -11/2 & 5/4 \\ 0 & 4 & -10 & 3 \\ 0 & 0 & 1/2 & -5/44 \\ 0 & 0 & 0 & 1/2 \end{pmatrix}.
\]
Now computing $Q^{-1}A$ we get

$$Q^{-1}A = \begin{pmatrix} 1 & 0 & -11/2 & 5/4 \\ 0 & 4 & -10 & 3 \\ 0 & 0 & 1/2 & -5/44 \\ 0 & 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 2 & 4 & 1 & 6 \\ 0 & 2 & 5 & 1 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 4 & -21 & 11/2 \\ 0 & 8 & -20 & 6 \\ 0 & 0 & 2 & 1/22 \\ 0 & 0 & 0 & 2 \end{pmatrix}. $$

And finally multiplying this last result on the right by $Q$ yields

$$(Q^{-1}A)Q = \begin{pmatrix} 2 & 4 & -21 & 11/2 \\ 0 & 8 & -20 & 6 \\ 0 & 0 & 2 & 1/22 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 11 & 0 \\ 0 & 1/4 & 5 & -4/11 \\ 0 & 0 & 2 & 5/11 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix},$$

which is the Jordan canonical form of the matrix $A$.

From this example, and from the structure of the Jordan Canonical Form, we see that, for each Jordan block, we are using certain homogeneous and non-homogeneous algebraic equations to compute the columns of the matrix $Q$ that defines the similarity transformation that puts the given matrix into
the Jordan form. Recall that, in the description of the Jordan canonical form, each Jordan block $J_i$ is and $n_i \times n_i$ block. Then, if $q_j$ is a column of $Q$ which involves the block $J_i$ then $n_1 + n_2 + \ldots + n_{i-1} + 1 \leq j \leq n_1 + n_2 + \ldots + n_i$. We can denote these columns of $Q$ by $q_{i,1}, q_{i,2}, \ldots, q_{i,n_i}$, so that

$$v_{i,j} = q_{n_1 + \ldots + n_{i-1} + j}.$$

It follows from the equation $AQ = QJ$ that

$$(*) \quad Aq_{i,1} = \lambda_i q_{i,1}, \quad Aq_{i,j+1} = \lambda_i q_{i,j+1} + q_{i,j}, \quad j = 1, 2, \ldots, n_i - 1.$$

The last equations above can actually be used to compute the matrix $Q$ that is required to put $A$ in Jordan canonical form: first compute the eigenvectors $q_{i,1}$ and then the remaining vectors $q_{i,j}$ for $j > 1$. It is important for future work to note that the series of equations for the columns of $Q$ can be written as a series of homogeneous equations. For example, if a Jordan block corresponding to a given $\lambda$ was a $3 \times 3$ block then the equations for the corresponding columns of $Q$ are

$$(A - \lambda I)q_1 = 0$$

$$(A - \lambda I)q_2 = q_1$$

$$(A - \lambda I)q_3 = q_2$$

But then multiplying both sides of the second equation $(A - \lambda I)q_2 = q_1$ by $(A - \lambda I)$ yields the equation

$$(A - \lambda I)^2 q_2 = (A - \lambda I)q_1 = 0$$

12
since $q_1$ is the eigenvector. Likewise, multiplication of the third equation yields another homogeneous equation $(A - \lambda I)^3 = O$ and so the set of equations for the eigenvectors and generalized eigenvectors become

$$
(A - \lambda I)q_1 = O
$$
$$
(A - \lambda I)^2q_2 = O
$$
$$
(A - \lambda I)^3, q_3 = O.
$$

The number of equations in this list, in this case $k$, depends on the size of the Jordan block. Of course we do not know that beforehand. However, given some matrix $A$ about which we have to a priori knowledge as to its Jordan form, or indeed about its eigenvalues and eigenvectors, we can proceed directly. Suppose, for the sake of argument, that we are presented with a $6 \times 6$ matrix. We first compute its eigenvalues; suppose that we find two of them, $\lambda_1 = \alpha$ of multiplicity 4 and $\lambda_2 = \beta$ of multiplicity 2. We then attempt to compute its eigenvectors.

In doing this, suppose that we find that there are two eigenvectors associated with $\lambda_1 = \alpha$, say $v_1$ and $v_2$ and only one eigenvector $v_3$ corresponding to $\lambda_2 = \beta$. Now, since $x_3$ has multiplicity 2 and there is only one eigenvector, we compute the generalized eigenvector by solving the equation

$$(A - \beta I)x = x_3.$$

We will then find a generalized eigenvector, say $x_4$, and, having done so, we have exhausted the possibilities for the eigenvalue $\lambda_2 = \beta$. The eigenvector $x_3$ and the generalized eigenvector will become the fifth and sixth columns $q_5$ and $q_6$ of the matrix $Q$ that we are attempting to construct.
We now take up the eigenvalue $\lambda_1 = \alpha$. Now this eigenvector has multiplicity 4 and we have found two eigenvectors $x_1$ and $x_2$ corresponding to it. At this stage, there are two possibilities for the Jordan form

$$
\begin{pmatrix}
\beta & 1 & 0 & 0 & 0 & 0 \\
0 & \beta & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha & 1 & 0 \\
0 & 0 & 0 & 0 & \alpha & 1 \\
0 & 0 & 0 & 0 & 0 & \alpha
\end{pmatrix}
\quad \text{or} \quad
\begin{pmatrix}
\beta & 1 & 0 & 0 & 0 & 0 \\
0 & \beta & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha & 1 & 0 & 0 \\
0 & 0 & 0 & \alpha & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha & 1 \\
0 & 0 & 0 & 0 & 0 & \alpha
\end{pmatrix}
$$

Note that both matrices has the required $2 \times 2$ block corresponding to $\beta$ and two eigenvectors corresponding to $\alpha$. The first matrix has a $3 \times 3$ block and a $1 \times 1$ block containing $\alpha$, while the second has two $2 \times 2$ blocks.

Remember that our goal is to find the columns of the matrix $Q$ and that its first column will be an eigenvector of $A$. So the best that we can say at this point is that $q_1$ must be a linear combination of the two eigenvectors that we have computed. Then a generalized eigenvector can be computed from the equation

$$(A - \alpha I) x = a x_1 + b x_2 := q_1.$$ 

but we do not know what values to choose for $a$ and $b$. However, as shown above, we can try to find a non-zero solution of $((A - \alpha I)^3 x = O$ and then a non-zero solutions of $(A - \alpha I)^2 x = O$. Note that since $(A - \alpha I)$ is singular, so are all its powers. Hence non-zero solutions will exist, provided the systems are consistent. If both systems are consistent, then there will be a three by three block corresponding to the eigenvalue $\alpha$, otherwise, two $2 \times 2$ blocks.
There is one fact that we need to check which is related to the chain of equations for the eigenvector and corresponding generalized eigenvectors. What we need to know is that the vectors computed in this way and used to construct the matrix $Q$ are, in fact, linearly independent. This means, of course, that $Q$ will be nonsingular.

**Theorem 2.1** Let $V$ be a vector space and $T : V \rightarrow V$ be linear. If $T$ is nilpotent of order $k^1$ for some integer $k \geq 1$ then the vectors $v, T(v), \ldots T^{k-1}(v)$ are linearly independent.

**Proof:** We will prove the result by induction. The result clearly holds in the case $k = 1$ since for any $v \in V$, we have $T(v) = 0$ whereas $T^0(v) = I(v) = v$.

Now suppose that the result holds for $k = n$. Then there is a vector $v$ satisfying

$$T^{n-1}(v) \neq 0 \text{ and } T^n(v) = 0.$$  

To show that the vectors are linearly independent, consider the equation

$$\alpha_1 v + \alpha_2 T(v) + \cdots + \alpha_n T^{n-1}(v) = 0.$$  

Now apply $T$ to both sides of the equation and use the fact that $T^n = O$.

This results in the equation

$$\alpha_1 T(v) + \alpha_2 T(v) + \cdots + \alpha_{n-1} T^{n-1}(v) = 0,$$

since $T(T^{n-1}) = T^n$. Set $w = T(v)$. Then the last line reads

$$\alpha_1 w + \alpha_2 T(w) + \cdots + \alpha_{n-2} T^{n-2}(w) = 0.$$  

---

1 Recall that $T$ is nilpotent of order $k$ provided $T^k(v) = 0$ for all $v \in V$ and $T^{k-1}(v_0) \neq 0$ for some $v_0 \in V$.  

15
By the induction hypothesis, the set
\[ \{w, T(w), \ldots, T^{n-2}(w)\} \]
is a linearly independent set and so \(\alpha_1 = \alpha_2 = \cdots = \alpha_n\). Hence the original linear combination reduces to \(\alpha_n T^{n-1} = 0\). It follows that \(\alpha_n = 0\) and so the vectors
\[ \{v, T(v), \ldots, T(v)\} \]
is a linearly independent set. This completes the proof.

**Exercise 2.2** Let \(L\) be a linear operator on a vector space \(V\) of dimension five and let \(A\) be any matrix representing \(L\). If \(L\) is nilpotent of order three, what are the possible Jordan canonical forms of \(A\)?

**Exercise 2.3** Let \(S\) be the subspace of \(C([a,b];\mathbb{R})\) spanned by \(x, xe^x\) and \(xe^x + x^2 e^x\). Let \(D\) be the differentiation operator on \(S\).

(a) Find a matrix \(A\) representing \(D\) with respect to the basis \(\{x, xe^x, xe^x + x^2 e^x\}\) in both the domain and range.

(b) Determine the Jordan canonical form of \(A\) and the corresponding basis for \(S\).

**Exercise 2.4** Find the Jordan canonical form, \(J\) for each of the following matrices and determine the matrix \(Q\) such that \(Q^{-1}AQ = J\)
(a) 

\[ A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 2 \\ 1 & -1 & 2 \end{pmatrix}. \]

(b) 

\[ A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \]