Lecture 22: Jordan canonical form of upper-triangular matrices (1)

Travis Schedler

Tue, Dec 6, 2011 (version: Tue, Dec 6, 1:00 PM)
Goals (2)

- Definition, existence, and uniqueness of Jordan canonical form of upper-triangular matrices
- How to compute Jordan form
- Relation with the minimal and characteristic polynomials

Read Chapters 8 and 10, do PS 11.
Warm-up: uniqueness of Jordan form (3)

(a) Let $N_n$ be the nilpotent, $n \times n$ upper-triangular matrix

$$N_n = \begin{pmatrix} 0 & 1 & & 0 \\ 0 & 1 & & \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \\ 0 & & 0 & 0 \end{pmatrix}.$$ 

Compute $\dim \ker(N_n^k)$ for all $k \geq 1$.

(b) Suppose that $A$ is block diagonal, with diagonal blocks $N_{n_i}$ for some $n_i \geq 1$. Using (a), show that

$$\dim(\ker(A^k)) - \dim(\ker(A^{k-1})) = \# \text{ blocks } N_{n_i} \text{ with } n_i \geq k,$$

where $\ker(A^0) := \ker(I) = \{0\}$.

(c) Conclude from (b) that, $\forall k$, $\# \text{ blocks } N_k$ is

$$(\dim(\ker(A^k)) - \dim(\ker(A^{k-1})))$$

$$- (\dim(\ker(A^{k+1})) - \dim(\ker(A^k))).$$
Solution to warm-up (4)

(a) The matrix $N^k_n$ is the same as $N_n$ but with the 1’s appearing a distance $k$ from the main diagonal (horizontally or vertically), rather than distance 1 (unless $k \geq n$). So the first $k$ columns of $N^k_n$ are zero, and the next $n - k$ are linearly independent. Hence $\dim \ker(N^k_n) = \min(k, n)$.

(b) $\dim(\ker(N^k_n)) - \dim(\ker(N^{k-1}_n)) = 1$ if $n \geq k$, and 0 if $n < k$. Now, $\dim(\ker(A^k)) = \text{sum of } \dim(\ker(N^k_{n_i}))$ over all diagonal blocks $N_{n_i}$. So $\dim(\ker(A^k)) - \dim(\ker(A^{k-1})) = \# \text{ of blocks } N_{n_i} \text{ with } n_i \geq k$.

(c) The number of blocks with $n_i \geq k$ minus the number of blocks with $n_i \geq k + 1$ is the number of blocks with $n_i = k$ exactly.
Theorem

If $T$ admits a basis in which $M(T)$ is upper-triangular, then it admits a basis in which $M(T)$ is block-diagonal with blocks

$$
\lambda I + N_n = \begin{pmatrix}
\lambda & 1 & 0 \\
\lambda & 1 & 0 \\
\lambda & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}.
$$

Such a matrix is called Jordan canonical form. It is unique for $T$ up to rearranging the order of the blocks.
Theorem
If $T$ admits a basis in which $\mathcal{M}(T)$ is upper-triangular, then it admits a basis in which $\mathcal{M}(T)$ is block-diagonal with blocks

$$
\lambda I + N_n = \begin{pmatrix}
\lambda & 1 & 0 \\
\lambda & 1 & \\
& & \ddots & \ddots \\
0 & & & 0 & 1 \\
0 & & & & \lambda
\end{pmatrix}.
$$

Such a matrix is called Jordan canonical form. It is unique for $T$ up to rearranging the order of the blocks.

Corollary: Over $\mathbb{C}$, every $T$ can be put in Jordan form.
Jordan canonical form (5)

Theorem

If \( T \) admits a basis in which \( \mathcal{M}(T) \) is upper-triangular, then it admits a basis in which \( \mathcal{M}(T) \) is block-diagonal with blocks

\[
\lambda I + N_n = \begin{pmatrix}
\lambda & 1 & \cdots & 0 \\
\lambda & 1 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & \lambda
\end{pmatrix}.
\]

Such a matrix is called Jordan canonical form. It is unique for \( T \) up to rearranging the order of the blocks.

Corollary: Over \( \mathbb{C} \), every \( T \) can be put in Jordan form.

Corollary: Over \( \mathbb{C} \), two matrices are conjugate iff they have the same Jordan canonical form (up to permuting blocks). More generally, this applies over any \( \mathbb{F} \), to matrices which are conjugate to upper-triangular ones.
Examples of and how to compute Jordan form (6)

- If $T$ admits an eigenbasis [orthonormal or not!] then its Jordan form is

$$
\begin{bmatrix}
0 & -1 \\
1 & 0
\end{bmatrix}
$$

- Upper-triangular examples, on the board! E.g.,

$$
\begin{bmatrix}
1 & 0 & 3 & 0 \\
0 & 1 & 5 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

- How to compute in general: For each eigenvalue $\lambda$, find $V(\lambda)$, compute $N = (T - \lambda I) | V(\lambda)$, find its Jordan form (e.g., compute $\text{dim null}(N^k)$ for all $k \geq 0$).
Examples of and how to compute Jordan form (6)

- If $T$ admits an eigenbasis [ortthonormal or not!] then its Jordan form is diagonal, and conversely.

- Upper-triangular examples, on the board! E.g.,

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

- How to compute in general: For each eigenvalue $\lambda$, find $V(\lambda)$, compute $N = (T - \lambda I)|_{V(\lambda)}$, find its Jordan form (e.g., compute $\text{dim null}(N^k)$ for all $k \geq 0$).
Examples of and how to compute Jordan form (6)

- If $T$ admits an eigenbasis [orthonormal or not!] then its Jordan form is diagonal, and conversely.

- If $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the rotation, its Jordan form

  $$
  \begin{bmatrix}
  0 & -1 \\
  1 & 0 
  \end{bmatrix}
  $$

- Upper-triangular examples, on the board! E.g.,

  $$
  \begin{bmatrix}
  1 & 0 & 3 & 0 \\
  0 & 1 & 5 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1 
  \end{bmatrix}
  $$

- How to compute in general: For each eigenvalue $\lambda$, find $V(\lambda)$, compute $N = (T - \lambda I) |_{V(\lambda)}$, find its Jordan form (e.g., compute $\dim \text{null}(N^k)$ for all $k \geq 0$).
Examples of and how to compute Jordan form (6)

- If $T$ admits an eigenbasis [orthonormal or not!] then its Jordan form is diagonal, and conversely.
- If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation, its Jordan form doesn’t exist; the corresponding complex matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has Jordan form $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.
Examples of and how to compute Jordan form (6)

▶ If $T$ admits an eigenbasis [orthonormal or not!] then its Jordan form is diagonal, and conversely.

▶ If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation, its Jordan form doesn’t exist; the corresponding complex matrix \[
\begin{pmatrix}
0 & -1 \\
1 & 0
\end{pmatrix}
\] has Jordan form
\[
\begin{pmatrix}
i & 0 \\
0 & -i
\end{pmatrix}.
\]

▶ Upper-triangular examples, on the board! E.g.,
\[
\begin{pmatrix}
1 & 0 & 3 & 0 \\
0 & 1 & 5 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]
Examples of and how to compute Jordan form (6)

- If $T$ admits an eigenbasis [orthonormal or not!] then its Jordan form is diagonal, and conversely.
- If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation, its Jordan form doesn’t exist; the corresponding complex matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has Jordan form $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.
- Upper-triangular examples, on the board! E.g.,
  $\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
- How to compute in general: For each eigenvalue $\lambda$, find $V(\lambda)$, compute $N = (T - \lambda I)|_{V(\lambda)}$, find its Jordan form (e.g., compute $\dim \ker(N^k)$ for all $k \geq 0$).
Fun applications of Jordan form (7)

- Important: compute functions of matrices: exp(A), sin(A), log(A), etc.
Fun applications of Jordan form (7)

- Important: compute functions of matrices: \( \exp(A), \sin(A), \log(A) \), etc.
- For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.
Fun applications of Jordan form (7)

- Important: compute functions of matrices: \( \exp(A), \sin(A), \log(A), \) etc.
- For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.
- Also, given a Taylor series at 0, e.g., \( 1 + x + \frac{x^2}{2!} + \cdots \), can plug in a nilpotent matrix and get a finite sum.
Fun applications of Jordan form (7)

- Important: compute functions of matrices: \( \exp(A), \sin(A), \log(A), \) etc.

- For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.

- Also, given a Taylor series at 0, e.g., \( 1 + x + x^2/2! + \cdots \), can plug in a nilpotent matrix and get a finite sum.

- For a Jordan nilpotent matrix, \( N^k \) is easy to write! Get immediately the answer.
Fun applications of Jordan form (7)

- Important: compute functions of matrices: \( \exp(A), \sin(A), \log(A), \) etc.
- For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.
- Also, given a Taylor series at 0, e.g., \( 1 + x + x^2/2! + \cdots \), can plug in a \textit{nilpotent} matrix and get a finite sum.
- For a Jordan nilpotent matrix, \( N^k \) is easy to write! Get immediately the answer.
- For \( \lambda I + N \), can still take \( \exp(\lambda I + N) = \exp(\lambda I) \exp(N) = e^\lambda \exp(N) \).
Fun applications of Jordan form (7)

- Important: compute functions of matrices: \( \exp(A), \sin(A), \log(A) \), etc.
- For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.
- Also, given a Taylor series at 0, e.g., \( 1 + x + x^2/2! + \cdots \), can plug in a nilpotent matrix and get a finite sum.
- For a Jordan nilpotent matrix, \( N^k \) is easy to write! Get immediately the answer.
- For \( \lambda I + N \), can still take \( \exp(\lambda I + N) = \exp(\lambda I) \exp(N) = e^\lambda \exp(N) \).
- For general \( f(x) \), on each \( V(\lambda) \), take Taylor series of \( f(x) \) at \( x = \lambda \) and plug in \( T|_{V(\lambda)} \): finite sum again, get \( f(T)|_{V(\lambda)} \).
Fun applications of Jordan form (7)

► Important: compute functions of matrices: \( \exp(A), \sin(A), \log(A) \), etc.

► For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.

► Also, given a Taylor series at 0, e.g., \( 1 + x + x^2/2! + \cdots \), can plug in a nilpotent matrix and get a finite sum.

► For a Jordan nilpotent matrix, \( N^k \) is easy to write! Get immediately the answer.

► For \( \lambda I + N \), can still take \( \exp(\lambda I + N) = \exp(\lambda I) \exp(N) = e^\lambda \exp(N) \).

► For general \( f(x) \), on each \( V(\lambda) \), take Taylor series of \( f(x) \) at \( x = \lambda \) and plug in \( T|_{V(\lambda)} \): finite sum again, get \( f(T)|_{V(\lambda)} \).

► We did this on PS 10 for \( f(x) = \sqrt{x} \) (Taylor series at \( x = 1 \))!
Fun applications of Jordan form (7)

- Important: compute functions of matrices: $\exp(A), \sin(A), \log(A)$, etc.
- For a diagonal matrix: Just take the function on each diagonal entry. For diagonalizable, can do it in a basis where the matrix is diagonal.
- Also, given a Taylor series at 0, e.g., $1 + x + x^2/2! + \cdots$, can plug in a nilpotent matrix and get a finite sum.
- For a Jordan nilpotent matrix, $N^k$ is easy to write! Get immediately the answer.
- For $\lambda I + N$, can still take $\exp(\lambda I + N) = \exp(\lambda I) \exp(N) = e^\lambda \exp(N)$.
- For general $f(x)$, on each $V(\lambda)$, take Taylor series of $f(x)$ at $x = \lambda$ and plug in $T|_{V(\lambda)}$: finite sum again, get $f(T)|_{V(\lambda)}$.
- We did this on PS 10 for $f(x) = \sqrt{x}$ (Taylor series at $x = 1$)!
- Works for $F = C$, and for $T$ which admit upper-tri matrices!
Solving differential equations (8)

- For example: a system $v'(x) = Av(x)$, for $A \in \text{Mat}(C, n, n)$, $v(x) \in \text{Mat}(C, n, 1)$ for all $x$ (function of $x$):

  \[
  v(x) = e^{Ax}. \text{ Now we know how to compute this!}
  \]

  - This also computes solutions to constant-coefficient ODEs $f^{(n)}(x) = a_0 f(x) + a_1 f'(x) + \cdots + a_{n-1} f^{(n-1)}(x)$: Write $v = (f, f', \ldots, f^{(n-1)})^t$ and then the system is the same as $v'(x) = Av(x)$, $A = \begin{pmatrix}
  0 & 0 & \cdots & 0 \\
  a_0 & 1 & 0 & \cdots & 0 \\
  1 & 0 & a_1 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  1 & 0 & a_{n-2} & \cdots & 1 \\
  0 & 1 & a_{n-1} & \cdots & 0
\end{pmatrix}.$
Solving differential equations (8)

- For example: a system $v'(x) = Av(x)$, for $A \in \text{Mat}(\mathbb{C}, n, n)$, $v(x) \in \text{Mat}(\mathbb{C}, n, 1)$ for all $x$ (function of $x$):
- Solution: $v(x) = e^{Ax}$. Now we know how to compute this! On each $V(\lambda)$, $e^{Ax}$ becomes $e^{(\lambda I + N)x} = e^{\lambda x}e^{Nx}$. 
Solving differential equations (8)

▷ For example: a system $\nu'(x) = A \nu(x)$, for $A \in \text{Mat}(\mathbb{C}, n, n)$, $\nu(x) \in \text{Mat}(\mathbb{C}, n, 1)$ for all $x$ (function of $x$):

▷ Solution: $\nu(x) = e^{Ax}$. Now we know how to compute this! On each $V(\lambda)$, $e^{Ax}$ becomes $e^{(\lambda I + N)x} = e^{\lambda x} e^{Nx}$.

▷ This also computes solutions to constant-coefficient ODEs $f^{(n)} = a_{0}f + a_{1}f' + \cdots + a_{n-1}f^{(n-1)}$: Write $\nu = (f, f', \ldots, f^{(n-1)})^t$ and then the system is the same as $\nu'(x) = A \nu(x)$,

$$
A = \begin{pmatrix}
0 & 0 & \cdots & 0 & a_{0} \\
1 & 0 & \cdots & 0 & a_{1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 1 & \cdots & 0 & a_{n-2} \\
0 & 0 & \cdots & 1 & a_{n-1}
\end{pmatrix}.
$$
Any Jordan matrix is obtained from bases of each of the \( V(\lambda) \). So, it suffices to assume there is only one eigenvalue \( \lambda \).

Up to replacing \( T \) with \( T - \lambda I \), we can assume \( T = N \) is nilpotent.

Now, we need only show that the sizes \( n_i \) that appear, and the number of times each appears, are independent of the basis.

This follows from the warm-up exercise! (The number of times \( n_i = k \) appears is 
\[
\dim \text{null}(A^{k+1}) - 2 \dim \text{null}(A^k) + \dim \text{null}(A^{k-1}).
\]
Proof of Jordan canonical form (10)

Jordan canonical form is based on the following key result. Let $N \in L(V)$ be nilpotent and for nonzero $v \in V$, let $m(v)$ be the maximum nonnegative integer such that $N^{m(v)}v \neq 0$. 

Lemma (Lemma 8.40)

There exist vectors $v_1, \ldots, v_k \in V$ such that $(v_1, Nv_1, \ldots, N^{m(v_1)}v_1, \ldots, v_k, Nv_k, \ldots, N^{m(v_k)}v_k)$ is a basis of $V$.

Note: the condition implies that $(N^{m(v_1)}v_1, \ldots, N^{m(v_k)}v_k)$ is a basis of null $N$ (consider any linear combination sent to zero by $N$).

▶ We prove by induction on $\dim V$. Since $\text{range } N \subsetneq V$, we can assume that $N|_{\text{range } N}$ has such vectors $u_1, \ldots, u_j$.

▶ Since $u_1, \ldots, u_j \in \text{range } N$, we can pick $v_1, \ldots, v_j$ with $Nv_i = u_i$.

▶ Furthermore, let $v_j+1, \ldots, v_k$ be vectors which extend $(N^{m(u_1)}v_1, \ldots, N^{m(u_j)}v_j)$ to a basis of null $N$. 

Proof of Jordan canonical form (10)

Jordan canonical form is based on the following key result. Let \( N \in \mathcal{L}(V) \) be nilpotent and for nonzero \( v \in V \), let \( m(v) \) be the maximum nonnegative integer such that \( N^{m(v)}v \neq 0 \).

**Lemma (Lemma 8.40)**

There exist vectors \( v_1, \ldots, v_k \in V \) such that \((v_1, Nv_1, \ldots, N^{m(v_1)}v_1, \ldots, v_k, Nv_k, \ldots, N^{m(v_k)}v_k)\) is a basis of \( V \).
Proof of Jordan canonical form (10)

Jordan canonical form is based on the following key result. Let $N \in \mathcal{L}(V)$ be nilpotent and for nonzero $v \in V$, let $m(v)$ be the maximum nonnegative integer such that $N^{m(v)}v \neq 0$.

**Lemma (Lemma 8.40)**

*There exist vectors $v_1, \ldots, v_k \in V$ such that $(v_1, Nv_1, \ldots, N^{m(v_1)}v_1, \ldots, v_k, Nv_k, \ldots, N^{m(v_k)}v_k)$ is a basis of $V$.*

Note: the condition implies that $(N^{m(v_1)}v_1, \ldots, N^{m(v_k)}v_k)$ is a basis of null $N$ (consider any linear combination sent to zero by $N$).
Proof of Jordan canonical form (10)

Jordan canonical form is based on the following key result. Let \( N \in \mathcal{L}(V) \) be nilpotent and for nonzero \( v \in V \), let \( m(v) \) be the maximum nonnegative integer such that \( N^{m(v)}v \neq 0 \).

**Lemma (Lemma 8.40)**

*There exist vectors \( v_1, \ldots, v_k \in V \) such that \((v_1, Nv_1, \ldots, N^{m(v_1)}v_1, \ldots, v_k, Nv_k, \ldots, N^{m(v_k)}v_k)\) is a basis of \( V \).*

Note: the condition implies that \((N^{m(v_1)}v_1, \ldots, N^{m(v_k)}v_k)\) is a basis of null \( N \) (consider any linear combination sent to zero by \( N \)).

- We prove by induction on \( \text{dim } V \). Since \( \text{range } N \subset V \), we can assume that \( N|_{\text{range } N} \) has such vectors \( u_1, \ldots, u_j \).
- Since \( u_1, \ldots, u_j \in \text{range } N \), we can pick \( v_1, \ldots, v_j \) with \( Nv_i = u_i \).
- Furthermore, let \( v_{j+1}, \ldots, v_k \) be vectors which extend \((N^{m(u_1)}u_1, \ldots, N^{m(u_j)}u_j)\) to a basis of null \( N \).
Completion of proof of the lemma (11)

- Claim: $v_1, \ldots, v_k$ give a basis of the desired form.
- We show linear independence. Suppose $\sum_{i,j} a_{i,j} N^j v_i = 0$. Applying $N$ yields $\sum_{i,j} a_{i,j} N^j u_i = 0$.
- By assumption that the $u_i$ give a basis of range $N$, we have $a_{i,j} = 0$ whenever $j \leq m(u_i) = m(v_i) - 1$.
- Thus the only nonzero coefficients are those of $(N^{m(v_1)} v_1, \ldots, N^{m(v_k)} v_k)$.
- These form a basis of null $N$. So all the coefficients are zero.
- Now, the length of the linearly independent list is $\dim \text{range } N + k = \dim \text{range } N + \dim \text{null } N = \dim V$, so it must be a basis. \(\square\)
In the reverse of the basis of the lemma, $\mathcal{M}(N)$ is block diagonal with $k$ blocks of sizes $m(v_k), \ldots, m(v_1)$, each of the form
\[
\begin{pmatrix}
0 & 1 \\
& & & \ddots \\
& & \ddots & 1 \\
& & & & 0
\end{pmatrix}.
\]
In the reverse of the basis of the lemma, \( \mathcal{M}(N) \) is block diagonal with \( k \) blocks of sizes \( m(v_k), \ldots, m(v_1) \), each of the form
\[
\begin{pmatrix}
0 & 1 \\
& \\
& \\
& \\
& & & & 1 \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & 0
\end{pmatrix}.
\]

Now, for a general operator \( T \), we can choose a basis as above for each \( V(\lambda) \), so that the nilpotent operator \( (T - \lambda I)|_{V(\lambda)} \) has the above form. Putting them together, \( \mathcal{M}(T) \) is in Jordan form.
A Jordan matrix is upper-triangular, so the char. poly. is
\[ \prod_{\lambda} (x - \lambda)^{d_{\lambda}}, \quad d_{\lambda} = \# \text{ of times } \lambda \text{ is on the diagonal}. \]
Char. and min. polys of Jordan matrices (13)

- A Jordan matrix is upper-triangular, so the char. poly. is $\prod_{\lambda} (x - \lambda)^{d_{\lambda}}$, $d_{\lambda} = \#$ of times $\lambda$ is on the diagonal.
- For each Jordan block $J = \lambda I + N_k$, $p(J) = 0$ if and only if $(x - \lambda)^k \mid p(x)$. 

Char. and min. polys of Jordan matrices (13)

- A Jordan matrix is upper-triangular, so the char. poly. is \( \prod_\lambda (x - \lambda)^{d_\lambda} \), \( d_\lambda = \# \) of times \( \lambda \) is on the diagonal.
- For each Jordan block \( J = \lambda I + N_k \), \( p(J) = 0 \) if and only if \( (x - \lambda)^k \mid p(x) \).
- So, the minimal polynomial is \( \prod_\lambda (x - \lambda)^{m_\lambda} \), where \( m_\lambda = \) the maximum size of Jordan block with eigenvalue (diag. entry) \( \lambda \).