Notes on the Jordan canonical form

Contents

1 Complex Jordan canonical form

Let A be a N-by-N matrix. The **Jordan canonical form** of A, J, is a matrix that verifies the following conditions:

- J is similar to A, i.e. there is an invertible matrix P such that $J = P^{-1}AP$.
- J is a diagonal block matrix of so-called **Jordan blocks**, which have the form

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{bmatrix}, \tag{1}$$

  where λ is an eigenvalue of A.

1.1 Most important theoretical results

Let $p(\lambda) = \Pi_{i=1}^{n}(\lambda_i - \lambda)^{m_i}$ with $\lambda_i \neq \lambda_j$ for $i \neq j$ and $m_i \in \mathbb{N}$ be the characteristic polynomial of the matrix A. The natural number $m_i = m_a(\lambda_i)$ is called the **algebraic multiplicity** of the eigenvalue λ_i.

The subspaces

$$E_i^j = E^j(\lambda_i) = \ker(A - \lambda_i I)^j, \tag{2}$$

where $j = \mathbb{N} \cup \{0\}$ play a major roll in the construction of the Jordan canonical form.

1.1.1 Result 1

Let $v \in E_i^j - E_i^{j-1}$, then the vectors defined iteratively by

$$\begin{aligned} v_j &= v & (3)\\ v_{j-1} &= (A - \lambda_i I)v_j & (4)\\ v_{j-2} &= (A - \lambda_i I)v_{j-1} & (5)\\ &\cdots \\ v_1 &= (A - \lambda_i I)v_2 & (6) \end{aligned}$$

are linearly independent, and $v_k \in E_i^k - E_i^{k-1}$ for $k \in \{1, \cdots, j\}$.

Furthermore, since

$$\begin{aligned} Av_1 &= \lambda_i v_1 & (7)\\ Av_k &= \lambda_i v_k + v_{k-1}\,, \quad k \in \{2, \cdots, j\}, & (8) \end{aligned}$$

the set $\{v_1, \cdots, v_j\}$ defines a Jordan block of size j associated to λ_i. These type of vectors are called **generalized eigenvectors**.

1.1.2 Result 2

Let $d_i^j = \dim(E_i^j)$ and $\bar{d}_i^j = d_i^j - d_i^{j-1}$. Then we have that d_i^j is an increasing sequence in j and that $\bar{d}_i^j$ is a decreasing sequence in j. In other words, the spaces E_i^j get larger and the "gaps" between get smaller, or they stay the same.

Either way, since these spaces are contained in a finite dimensional vector space, the chain of spaces $\{E_i^j\}_{j=1}^{\infty}$ has to stabilize after an index, ν_i. This index is actually called the **index** of the eigenvalue λ_i. Hence, we have that

$$\begin{aligned} \{0\} &= E_i^0 \subsetneq E_i^1 \subsetneq \cdots \subsetneq E_i^{\nu_i - 1} \subsetneq E_i^{\nu_i} & (9)\\ E_i^{\nu_i} &= E_i^{\nu_i+1} = E_i^{\nu_i+2} = \cdots & (10) \end{aligned}$$

1.1.3 Result 3

$$\mathbb{R}^N = \bigoplus_{i=1}^{m} E_i^{\nu_i}\,, \tag{11}$$

i.e. each vector $v \in \mathbb{R}^N$ has a unique representation as $v = \sum_{i=1}^m v_i$ with $v_i \in E_i^{\nu_i}$.

Moreover, the spaces $E_i^{\nu_i}$ are invariant under A. In other words, this allows us to find the generalized vectors and Jordan blocks associated to one eigenvalue at the time, and then concatenate them all together in order to find the Jordan canonical form, J and an associated similarity transformation matrix, P.

1.2 Algorithm

Results 1, 2 and 3 provide a method for finding a base $\mathcal{B}$ for $\mathbb{R}^N$ such that the matrix A represented with respect to this new base is its Jordan canonical form, J. In other words, we can choose the columns of the associated similarity transform matrix, P, as the vectors in $\mathcal{B}$.

This is achieved in the following way: For each eigenvalue λ_i of A, the spaces $\{E_i^j\}_{j=1}^{\nu_i}$ are calculated. Starting at $E_i^{\nu_i}$ and moving down to E_i^0, one checks whether there are $\bar{d}_i^j$ vectors in $\mathcal{B} \cap (E_i^j - E_i^{j-1})$. If not, enough linear independent vectors $\{v_{l,j}\}_{l=1}^{l_0} \in E_i^j - E_i^{j-1}$ are found so that the set $\{v_{l,j}\}_{l=1}^{l_0} \cup (\mathcal{B} \cap (E_i^j - E_i^{j-1}))$ contains $\bar{d}_i^j$ linearly independent vectors. All the new vectors $v_{l,j}$ are "back-propagated" as done in Result 1:

$$v_{l,k} = (A - \lambda_i I)v_{l,k+1} \quad k \in \{1, \cdots, j-1\} \tag{12}$$

Hence, each set $\{v_{l,1}, \cdots, v_{l,j}\}$ defines a Jordan block of size j associated to λ_i, and it is added to $\mathcal{B}$ without altering the sequence order. This is done for all $l \in \{1, \cdots, l_0\}$. Finally, we move from E_i^j to E_i^{j-1}.

On the other hand, if there are $\bar{d}_i^j$ vectors in $\mathcal{B} \cap (E_i^j - E_i^{j-1})$, we just move from E_i^j to E_i^{j-1}. We stop when we arrive at E_i^0 because we then have found all Jordan blocks associated to λ_i. We repeat this process with the next eigenvalue λ_{i+1}, if any is left.

This algorithm is summarized as follows:

Initialization: $\mathcal{B} \leftarrow \emptyset$;
for $i = 1 \rightarrow m$ **do**
  for $j = \nu_i \rightarrow 1$ **do**
    $l_0 \leftarrow$ difference between $\bar{d}_i^j$ and the number of elements in $\mathcal{B} \cap (E_i^j - E_i^{j-1})$;
    if $l_0 \geq 1$ **then**
      find vectors $\{v_{l,j}\}_{l=1}^{l_0} \in E_i^j - E_i^{j-1}$ such that $\{v_{l,j}\}_{l=1}^{l_0} \cup (\mathcal{B} \cap (E_i^j - E_i^{j-1}))$ contains $\bar{d}_i^j$ linearly independent vectors;
      for $l = 1 \rightarrow l_0$ **do**
        for $k = j-1 \rightarrow 1$ **do**
          $v_{l,k} = (A - \lambda_i I)v_{l,k+1}$
        end
        add a Jordan block of size j associated to λ_i to J;
        $\mathcal{B} \leftarrow \mathcal{B} \cup \{v_{l,1}, \cdots, v_{l,j}\}$
      end
    end
  end
end

NB: The number of Jordan blocks of size j associated to λ_i is equal to $\bar{d}_i^j - \bar{d}_i^{j+1}$. Hence, the Jordan canonical form can be found just by calculating the dimensions of the spaces E_i^j.

NB: The Jordan canonical form is unique up to permutations of Jordan blocks.

1.3 Example

Consider the matrix

$$A = \begin{bmatrix} 2 & -4 & -2 & 0 & -8 & 0 & 12 \\ 1 & 6 & 1 & 0 & 4 & 0 & -6 \\ 0 & 0 & 4 & 0 & 4 & 0 & -8 \\ 0 & 0 & -1 & 4 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 12 & 0 & 8 & 8 & 0 & -8 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \end{bmatrix} \tag{13}$$

Its characteristic polynomial is $p(\lambda) = -\lambda^3(4-\lambda)^4$. Moreover,

$$E^1(0) = \ker A = \operatorname{span}\{2e_1 - e_2 - 2e_3 - e_7, e_6\} \tag{14}$$
$$E^2(0) = \ker A^2 = \operatorname{span}\{2e_1 - e_2 + 2e_5, e_6, 2e_3 + 2e_5 + e_7\} \tag{15}$$
$$E^1(4) = \ker(A - 4I) = \operatorname{span}\{2e_1 - e_2 + e_4, e_4 + 2e_6\} \tag{16}$$
$$E^2(4) = \ker(A - 4I) = \operatorname{span}\{e_1, e_2 - e_4, e_3, e_4 + 2e_6\} \tag{17}$$

Since $\dim(E^2(0)) = 3 = m_a(0)$, it follows that the index of the eigenvalue 0 is 2. Analogously, we have that the index of the eigenvalue 4 is 2.

The basis $\mathcal{B} = \{v_i\}_{i=1}^7$ is constructed in the following way:

$$v_2 = 2e_1 - e_2 + 2e_5 \qquad \in E^2(0) - E^1(0) \tag{18}$$
$$v_1 = Av_1 = -8e_1 + 4e_2 + 8e_3 + 8e_6 + 4e_7 \qquad \in E^1(0) \tag{19}$$
$$v_3 = e_6 \qquad \in E^1(0) \tag{20}$$
$$v_5 = e_1 \qquad \in E^2(4) - E^1(4) \tag{21}$$
$$v_4 = (A - 4I)v_5 = -2e_1 + e_2 + 2e_6 \qquad \in E^1(4) \tag{22}$$
$$v_7 = e_3 \qquad \in E^2(4) - E^1(4) \tag{23}$$
$$v_6 = (A - 4I)v_7 = -2e_1 + e_2 - e_4 \qquad \in E^1(4) \tag{24}$$

Hence,

$$P = \begin{bmatrix} -8 & 2 & 0 & -2 & 1 & -2 & 0 \\ 4 & -1 & 0 & 1 & 0 & 1 & 0 \\ 8 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 1 & 2 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} \tag{25}$$

2 Real Jordan canonical form

If A is in addition a real matrix, then $v \in E^j(\lambda)$ if and only if $\bar{v} \in E^j(\bar{\lambda})$.

The real Jordan canonical form is constructed by modifying the complex Jordan canonical in the following way:

For each non-real eigenvalue λ:

- Eliminate the Jordan blocks in J associated to $\bar{\lambda}$ and the columns in P corresponding to generalized eigenvectors associated to $\bar{\lambda}$.
- For each Jordan block in J associated to λ, replace all elements in the diagonal by the 2-by-2 matrix
$$\begin{bmatrix} \Re(\lambda) & \Im(\lambda) \\ -\Im(\lambda) & \Re(\lambda) \end{bmatrix}. \tag{26}$$
Moreover, replace all other elements in the Jordan block, which are 0's or 1's, by a 2-by-2 zero or identity matrix, respectively.
- Replace each column in P corresponding to a generalized eigenvector associated to λ with two columns: the real and imaginary part of the original column, in that order.

3 Tips and tricks

The most effective method for finding the kernel of a matrix, and therefore the most effective method for finding generalized eigenvectors, is Gaussian column elimination.

For example, in order to find $E^1(\lambda)$, one starts with $[A - \lambda I|I]$ and performs column operations until the maximum number of zero columns on the left side of the augmented matrix is achieved. In symbols,

$$[A - \lambda I|I] \sim [(A - \lambda I)C|C], \tag{27}$$

where C is an invertible matrix.

Another important observation is that if one has used Gaussian column elimination to find a basis for $E^j(\lambda)$, one can reuse all the work done to calculate a basis for $E^{j+1}(\lambda)$. In symbols,

$$[(A - \lambda I)^j|I] \sim [(A - \lambda I)^j C|C] \tag{28}$$
$$\Rightarrow [(A - \lambda I)^{j+1}|I] \sim [(A - \lambda I)^{j+1} C|C]. \tag{29}$$

After calculating $E^j(\lambda)$, $(A - \lambda I)^j C$ is known, and this matrix has usually more zeros than $(A - \lambda I)^j$. Hence, the calculation of $(A - \lambda I)^{j+1} C = (A - \lambda I)(A - \lambda I)^j C$ can be done relatively fast, and the resulting matrix has usually even more zeros than $(A - \lambda I)^j C$.