How to compute the Jordan form of a matrix

§1. Notation.

(1.1) Let $A \in \mathrm{M}_n(\mathbb{C})$ be a square matrix with entries in $\mathbb{C}$. Let

$$f(x) := \det(x \cdot \mathrm{I}_n - A)$$

be the characteristic polynomial. Suppose that $f(x)$ has already been factored. (So we have swept the most difficult part "under the rug".) Write

$$f(x) = \prod_{i=1}^{r} (x-\lambda_i)^{d_i} = (x-\lambda_1)^{d_1} \cdots (x-\lambda_r)^{d_r},$$

where $\lambda_1, \ldots, \lambda_r$ are mutually distinct complex numbers, and $d_1, \ldots, d_r \in \mathbb{N}_{\geq 1}$ are positive integers. In other words $\lambda_1, \ldots, \lambda_r$ are the eigenvalues of A, and $d_1, \ldots, d_r$ are their multiplicities as roots of $f(x)$.

(1.2) Computing the Jordan form of A means that you want to find a good basis of $\mathbb{C}^n_{\mathrm{col}}$, grouped in blocks

$$v_{1,1}, v_{1,2}, \ldots, v_{1,e_1}; \ldots, v_{j,1}, v_{j,2}, \ldots, v_{j,e_j}, \ldots, v_{s,1}, \ldots, v_{s,e_s}.$$

Here s is the number of (Jordan) blocks, and $e_1, \ldots, e_s$ are the sizes of the Jordan blocks. Moreover there is one of the eigenvalues associated to each of the s blocks; we will use the notation $\lambda_{c(1)}, \ldots, \lambda_{c(s)}$ for the eigenvalue associated with the s blocks. We have

$$\sum_{1 \leq j \leq s,\ c(j)=i} e_j = d_i \qquad \forall i = 1, \ldots, r$$

In other words, the sum of the sizes of all Jordan blocks with eigenvalue λ_i is the multiplicity of λ_i in the characteristic polynomial. The theory of Jordan canonical forms asserts the existence of a good basis vectors as above, which satisfies the properties to be reviewed in 1.3 below.

(1.3) For each block of the good basis vectors $v_{j,1}, v_{j,2}, \ldots, v_{j,e_j}$, we have

$$A \cdot v_{j,1} = \lambda_{c(j)} v_{j_1},\ \ A \cdot v_{j,2} = \lambda_{c(j)} v_{j_2} + v_{j,1}, \cdots,\ \ A \cdot v_{j,e_j} = \lambda_{c(j)} v_{j,e_j} + v_{j,e_{j-1}}$$

Let C be the invertible $n \times n$ matrix whose columns are the good basis vectors

$$v_{1,1}, v_{1,2}, \ldots, v_{1,e_1}; \ldots, v_{j,1}, v_{j,2}, \ldots, v_{j,e_j}, \ldots, v_{s,1}, \ldots, v_{s,e_s}.$$

Then we the matrix $C^{-1} \cdot A \cdot C$ takes a simple form, call the Jordan form of A. It is composed of s diagonal Jordan blocks; the j-th Jordan block J_j has size e_j, with associated eigenvalue $\lambda_{c(j)}$:

$$J_j = \lambda_{c(j)} \cdot \mathrm{I}_{e_j} + N_{e_j},$$

where $N_{e_j} \in \mathrm{M}_{e_j}(\mathbb{C})$ is the strictly upper triangular $e_j \times e_j$ matrix such that the the $e_j - 1$ entries immediately above the diagonal are 1, and all other entries are 0.

The collection of pairs

$$\left\{ (e_1, \lambda_{c(1)}), \ldots, (e_s, \lambda_{c(s)}) \right\},$$

which specifies the Jordan form of A, is determined by A up to permutation of the s pairs.

(1.4) Computing the Jordan form of A means finding a good basis vector

$$v_{1,1}, v_{1,2}, \ldots, v_{1,e_1}; \ldots, v_{j,1}, v_{j,2}, \ldots, v_{j,e_j}, \ldots, v_{s,1}, \ldots, v_{s,e_s}$$

grouped in blocks, satisfying the properties in 1.3. Textbooks in linear algebra tends to end with the proof of the existence of Jordan canonical form (and rational canonical forms). Systematic ways to to actually compute a good basis tend not to make it to the actual text. However it is not too difficult to produce an algorithm. We present one in this note.

(1.5) REMARK Let's note that it is quite straight-forward to compute the generalized eigenspaces of A, from which one can easily compute the matrix exponential $\exp(tA)$ without finding a good basis which turns A into a Jordan form.

Recall that for each eigenvalue λ_i, the linear span of all blocks of the good basis with eigenvalue λ_i is the *generalized eigenspace* for the eigenvalue λ_i; we will write it as $V(\lambda_i)$. We have

$$\dim(V(\lambda_i)) = \sum_{1\leq j\leq s,\ c(j)=i} e_j = d_i.$$

The generalized eigenspace $V(\lambda_i)$ can be computed easily:

$$V(\lambda_i) = \mathrm{Ker}((A-\lambda_i)^{d_i}).$$

§2. How to compute the invariants

(2.1) The problem of computing a good basis of $\mathbb{C}$ for A can be dealt with one eigenvalue at a time. Let's fix one of the eigenvalues λ_{i_0}, and abbreviate it to λ. Changing the natation slightly, our goal is to find a good basis

$$w_{1,1}, \ldots, w_{1,e_1}; \ldots; w_{t,e_t}$$

of the generalized eigenspace $V(\lambda)$ for the eigenvalue λ. (They are members of the good basis vectors of $\mathbb{C}^n$ associated to the eigenvalue λ.) Let $d = d(\lambda) = e_1 + \cdots + e_t =$ multiplicity of λ as a root of the characteristic polynomial $f(x)$. We may and do assume that $e_1 \leq e_2 \cdots \leq e_t$.

(2.2) The positive integers $e_1, \ldots, e_t$ are the sizes of Jordan blocks associated to the eigenvalue λ. Their sum is the multiplicity of λ. They can be easily determined from

$$k_h := \dim\left(\mathrm{Ker}(A-\lambda\cdot \mathrm{I}_n)^h\right), \quad h = 1, \ldots, d.$$

The numbers k_j's satisfy

$$1 \leq k_1 \leq k_2 \leq \cdots \leq k_{d-1} \leq k_d = d.$$

They are related to the yet-to-be-computed numbers t and $e_1, \ldots, e_t$ by

$$k_h = \sum_{j=1}^{t} \mathrm{Min}(h, e_j).$$

Define natural numbers $e_1, \ldots, e_d$ by

$$\ell_1 := k_1,\ \ell_2 = k_2 - k_1,\ \ell_3 = k_3 - k_2,\ \cdots,\ \ell_d = k_d - k_{d-1}.$$

For each $h = 1, \ldots, d$,

$$\ell_h = \text{number of Jordan blocks for } \lambda \text{ of size } \geq h,$$

so we have

$$k_1 = \ell_1 \geq \ell_2 \geq \cdots \geq \ell_d.$$

Define natural number $m_1, \ldots, m_d$ by

$$m_1 := \ell_1 - \ell_2,\ m_2 := \ell_2 - \ell_3,\ \ldots,\ m_{d-1} = \ell_{d-1} - \ell_d,\ m_d := \ell_d.$$

Then for each $h = 1, \ldots, d$, we have

$$m_h = \text{number of Jordan blocks for } \lambda \text{ of size } h.$$

In other words, if we start with the sequence $m_1, m_2, \ldots, m_d$, replace the entry m_h by the sequence $h, \ldots, h$ of length m_h if $m_h \geq 1$, and eliminate all entries of 0, we get the sizes $e_1, \ldots, e_t$ of Jordan blocks associated to the eigenvalue λ.

§3. How to compute a good basis

(3.1) We will use the notation in 2.1 and fix an eigenvalue λ, with multiplicity d in the characteristic polynomial $f(x)$ of A. In 2.2 we have defined a sequence of natural numbers $m_1, \ldots, m_d$, which are easily computed from the dimensions of $\mathrm{Ker}((A-\lambda)^h)$.

Let

$$q(x) := \frac{f(x)}{(x-\lambda)^d},$$

a polynomial of degree $n-d$. The image $q(A)(\mathbb{C}^n)$ is the generalized eigenenspace $V(\lambda)$, i.e.

$$q(A)(\mathbb{C}^n) = \mathrm{Ker}((A - \lambda \cdot \mathrm{I}_n)^d) =: V(\lambda).$$

Define several vector subspaces of $V(\lambda)$.

1. $U := (A - \lambda \cdot \mathrm{I}_n)(q(A)(\mathbb{C}^n)) = (A - \lambda \cdot \mathrm{I}_n)(V(\lambda))$.
2. $V_h := \mathrm{Ker}((A-\lambda)^h)$ for $h = 1, \ldots, d$.
3. $W_h := U + V_h$ for $h = 1, \ldots, d$

Clearly

$$W_0 := U \subseteq W_1 \subseteq \cdots \subseteq W_d.$$

The existence of Jordan canonical forms tells us that the total number t of Jordan blocks associated to the eigenvalue λ is

$$t = d - \dim(U)$$

and

$$w_h > 0 \quad \text{if and only if} \quad W_h \supsetneqq W_{h-1}$$

for all $h = 1, \ldots, h$.

(3.2) The algorithm

Step 1. Compute the sequence of subspaces

$$W_0 \subseteq W_1 \subseteq \cdots \subseteq W_d$$

of $V(\lambda)$. More precisely, compute a $\mathbb{C}$-basis

$$u_1, \ldots, u_{d-t}, w_1, w_2, \ldots, w_t$$

adapted to the increasing sequence of subspaces

$$W_0 \subseteq W_1 \subseteq \cdots \subseteq W_d$$

in the following sense

(i) For each h with $0 \leq h \leq d$, there exists a unique natural number a with $0 \leq a \leq t$ such that

$$W_h = \mathbb{C}u_1 + \cdots + \mathbb{C}u_{d-t} + \mathbb{C}w_1 + \cdots + \mathbb{C}w_a.$$

(ii) If $W_{h-1} \subsetneqq W_h$, write

$$W_h = \mathbb{C}u_1 + \cdots + \mathbb{C}u_{d-t} + \mathbb{C}w_1 + \cdots + \mathbb{C}w_a, \quad W_{h-1} = \mathbb{C}u_1 + \cdots + \mathbb{C}u_{d-t} + \mathbb{C}w_1 + \cdots \mathbb{C}w_b$$

with $b < a$, then

$$w_{b+1}, \ldots, \ldots w_a \in V_h.$$

Step 2. For each $j = 1, \ldots, t$, let

$$e_j := \mathrm{Min}\{k \mid w_j \in V_k\}.$$

(These numbers $e_1, \ldots, e_t$ are easily read off from the computation in step 1.) Define

$$v_{j,1} := (A - \lambda \cdot \mathrm{I}_n)^{e_j - 1} w_j, \;\; v_{j,2} := (A - \lambda \cdot \mathrm{I}_n)^{e_j - 1} w_j, \;\; \ldots, \;\; v_{j,e_j - 1} := (A - \lambda \cdot \mathrm{I}_n) w_j, \;\; v_{j,e_j} := w_j$$

for $j = 1, \ldots, t$. Thus we get a block of good basis vectors for $V(\lambda)$ for each $j = 1, \ldots, t$. Putting the t blocks together we obtain a good basis for the generalized eigenspace $V(\lambda)$.

(3.3) **Remark** The algorithm above uses only Gaussian eliminations. The actual computation takes places in step 1.