The Jordan canonical form

Francisco–Javier Sayas
University of Delaware

November 22, 2013

The contents of these notes have been translated and slightly modified from a previous version in Spanish. Part of the notation for the iterated kernels and many of the ideas for the proof of Jordan's Decomposition Theorems are borrowed from [1]. The flavor to functional analysis (Riesz theory) is taken from [2]. Finally, notation has been adapted from [3] to fit in what I was teaching in MATH 672 in the Fall 2013 semester.

1 Preliminary notions

Definition 1. *Let $A \in \mathcal{M}(n;\mathbb{F})$, and let $\mathcal{V}$ be a subspace of $\mathbb{F}^n = \mathbb{F}^n_c$. We say that $\mathcal{V}$ is $A-$invariant if*

$$A\,u \in \mathcal{V}, \qquad \forall u \in \mathcal{V}.$$

Examples. If λ is an eigenvalue of A, then $\ker(A-\lambda I) = \{u \in \mathbb{F}^n \,:\, A\,u = \lambda\,u\}$ is $A-$invariant. More generally

$$\ker(A-\lambda I)^j \qquad \text{is } A-\text{invariant} \qquad \forall j.$$

Proof. If $u \in \ker(A-\lambda I)^j$, then

$$\begin{aligned}(A-\lambda\,I)^j u = 0 \qquad &\Longrightarrow \qquad 0 = A\,(A-\lambda\,I)^j u \overset{(*)}{=} (A-\lambda\,I)^j A\,u \\ &\Longrightarrow \qquad Au \in \ker(A-\lambda I)^j.\end{aligned}$$

To prove (*), we just have to notice that

$$(A-\lambda\,I)^j = \sum_{i=0}^{j} (-\lambda)^{j-i} \binom{j}{i} A^i,$$

and therefore $A\,(A-\lambda\,I)^j = (A-\lambda\,I)^j\,A$. □

Definition 2. *If $A \in \mathcal{M}(n;\mathbb{F})$ and $\mathbb{F}^n = \mathcal{V}\oplus\mathcal{W}$, where $\mathcal{V}$ and $\mathcal{W}$ are $A-$invariant, we say that $\mathbb{F}^n = \mathcal{V}\oplus\mathcal{W}$ is an $A-$invariant decomposition. The same definition can be extended to*

$$\mathbb{F}^n = \mathcal{V}_1 \oplus \ldots \oplus \mathcal{V}_k,$$

where $\mathcal{V}_j$ is $A-$invariant for all j.

Example. If A is diagonalizable and $\lambda_1, \dots, \lambda_k$ are the distinct eigenvalues of A, then

$$\mathbb{F}^n = \ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_k I)$$

is an $A-$invariant decomposition. Also if

$$Av_j = \mu_j v_j, \qquad \forall j \qquad \text{and } \{v_1, \dots, v_n\} \text{ is a basis for } \mathbb{F}^n,$$

then

$$\mathbb{F}^n = \operatorname{span}[v_1] \oplus \operatorname{span}[v_2] \oplus \dots \oplus \operatorname{span}[v_n]$$

is an $A-$invariant decomposition.

The associated matrix. If $\mathbb{F}^n = \mathcal{V} \oplus \mathcal{W}$ is an $A-$invariant decomposition, and we construct a basis for $\mathbb{F}^n$ as follows:

$$\underbrace{v_1, \dots, v_\ell}_{\text{basis for } \mathcal{V}}, \underbrace{v_{\ell+1}, \dots, v_n}_{\text{basis for } \mathcal{W}}$$

and we build the matrix

$$P = \left[\begin{array}{c|c|c|c|c|c} v_1 & \dots & v_\ell & v_{\ell+1} & \dots & v_n \end{array} \right],$$

then

$$P^{-1} A P = \begin{bmatrix} A_{\ell \times \ell} & 0 \\ 0 & A_{n-\ell \times n-\ell} \end{bmatrix}$$

In general, invariant decomposition produce **block diagonal matrices**. Note that if

$$P^{-1} A P = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

where A_j is a square matrix for all j, then: (a) the corresponding partition of the columns of P creates an $A-$invariant decomposition of $\mathbb{F}^n$ in k subspaces; (b) the characteristic polynomial of A is the product of the characteristic polynomials of the matrices A_j.

First goal. In a first step we will try to find a matrix P such that

$$P^{-1} A P = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

where

$$\chi_{A_j}(x) = (\lambda_j - x)^{m_j} \qquad \lambda_i \neq \lambda_j \qquad \forall i \neq j.$$

This will be possible for any matrix such that

$$\chi_A(x) = (\lambda_1 - x)^{m_1} \dots (\lambda_k - x)^{m_k} \qquad m_1 + \dots + m_k = n.$$

We will have thus created an $A-$invariant decomposition

$$\mathbb{F}^n = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_k,$$

where $\dim \mathcal{V}_j = m_j$. *From now on, we will restrict our attention to the case* $\mathbb{F} = \mathbb{C}$. The case where A is a real matrix whose characteristic polynomial has only real roots will follow as a particular case.

2 Iterated kernels

Notation. From now on we will consider a matrix whose characteristic polynomial is

$$\chi_A(x) = (\lambda_1 - x)^{m_1} \dots (\lambda_k - x)^{m_k},$$

with pairwise different roots $\{\lambda_1, \dots, \lambda_k\}$ with multiplicities $m_j \geq 1$.

Definition 3. *Let* $A \in \mathcal{M}(n; \mathbb{C})$ *and* λ *be an eigenvalue of* A. *The subspaces*

$$E_j(\lambda) = \ker(A - \lambda\, I)^j, \qquad j \geq 1,$$

are called iterated kernels. We denote $E_0(\lambda) = \ker(\, I\,) = 0$.

Elementary properties.

(1) $E_1(\lambda)$ is the set of all eigenvectors associated to λ and the zero vector (which is never considered an eigenvector).

(2) $E_j(\lambda) \subseteq E_{j+1}(\lambda)$ for all j.

Proof. It follows from the following straightforward argument

$$(A - \lambda\, I)^j u = 0 \qquad \Longrightarrow \qquad (A - \lambda\, I)^{j+1} u = 0.$$

□

(3) If $E_\ell(\lambda) = E_{\ell+1}(\lambda)$, then $E_{\ell+1}(\lambda) = E_{\ell+2}(\lambda)$. Consequently

$$E_\ell(\lambda) = E_{\ell+1}(\lambda) = E_{\ell+2}(\lambda) = \dots.$$

Proof. Let u be such that $(A - \lambda\, I)^{\ell+2} u = 0$ and define $v = (A - \lambda\, I)u$. Then

$$(A - \lambda I)^{\ell+1} v = (A - \lambda I)^{\ell+2} u = 0,$$

which implies that

$$v \in E_{\ell+2}(\lambda) = E_{\ell+1}(\lambda),$$

and therefore

$$(A - \lambda I)^{\ell+1} u = (A - \lambda I)^{\ell} v = 0.$$

This argument shows then that $E_{\ell+2}(\lambda) \subseteq E_{\ell+1}(\lambda) \subseteq E_{\ell+2}(\lambda)$. □

(4) The spaces $E_j(\lambda)$ are $A-$invariant.

Not so elementary properties. The proofs for the following results will be given in Section 3.

(5) For all $j \geq 1$

$$\dim E_{j+1}(\lambda) - \dim E_j(\lambda) \leq \dim E_j(\lambda) - \dim E_{j-1}(\lambda).$$

(6) Let $\lambda_1, \ldots, \lambda_k$ be eigenvalues of A. Let $j_1, \ldots, j_k \geq 1$. Then, the following sum of subspaces is direct:

$$E_{j_1}(\lambda_1) \oplus E_{j_2}(\lambda_2) \oplus \ldots \oplus E_{j_k}(\lambda_k).$$

In other words, the spaces $E_{j_1}(\lambda_1), \ldots, E_{j_k}(\lambda_k)$ are independent.

(7) The maximum dimension of the iterated kernels $E_k(\lambda)$ is the multiplicity of λ as a root of the characteristic polynomial. In other words, for every eigenvalue λ, there exists ℓ such that

$$E_\ell(\lambda) = E_{\ell+1}(\lambda), \qquad \dim E_\ell(\lambda) = m_\lambda,$$

where m_λ is the algebraic multiplicity of λ as a root of χ_A.

On the dimensions of the iterated kernels. Let

$$n_j = \dim E_j(\lambda) = \dim \ker(A - \lambda I)^j, \qquad j \geq 0$$

(note that $n_0 = 0$). The preceding properties imply that this sequence of integers satisfies the following properties:

$$
\begin{aligned}
&(1) \qquad 0 = n_0 < n_1 < \ldots < n_\ell = n_{\ell+1} = \ldots,\\
&(2) \qquad n_\ell - n_{\ell-1} \leq n_{\ell-1} - n_{\ell-2} \leq \ldots \leq n_2 - n_1 \leq n_1 - n_0 = n_1\\
&(3) \qquad n_\ell = m_\lambda.
\end{aligned}
$$

The sequence of iterated kernels is then

$$
\begin{array}{lccccccccccccc}
 & E_0(\lambda) & \subset & E_1(\lambda) & \subset & E_2(\lambda) & \subset & \ldots & \subset & E_\ell(\lambda) & = & E_{\ell+1}(\lambda) & = & \ldots\\
\dim: & 0 = n_0 & < & n_1 & < & n_2 & < & \ldots & < & n_\ell & = & n_{\ell+1} & = & \ldots
\end{array}
$$

A simple but important observation. If $B = P^{-1}AP$, then

$$u \in \ker(B - \lambda I)^j \qquad \Longleftrightarrow \qquad P\,u \in \ker(A - \lambda I)^j.$$

Therefore, the dimensions of the iterated kernels are invariant by similarity transformations. This can be also explained in different words: the dimensions of the iterated kernels depend on the operator and not on the particular matrix representation of the operator.

Theorem 1 (First Jordan decomposition theorem)**.** *If*

$$\chi_A(x) = (\lambda_1 - x)^{m_1}(\lambda_2 - x)^{m_2}\dots(\lambda_k - x)^{m_k},$$

then there exist integers ℓ_j such that

$$\mathbb{C}^n = E_{\ell_1}(\lambda_1) \oplus \dots \oplus E_{\ell_k}(\lambda_k)$$

and there is a change of basis such that

$$P^{-1}\,A\,P = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix},$$

where

$$\chi_{A_p}(x) = (\lambda_p - x)^{m_p} \qquad \forall p.$$

Furthermore, the dimensions of the iterated kernels of the submatrices A_p coincide with the dimensions of the kernels $E_j(\lambda_p)$ for the matrix A, that is,

$$\dim\ker(A_p - \lambda_p I)^j = \dim\ker(A - \lambda_p I)^j \qquad \forall j.$$

Proof. The first assertion is a direct consequence of properties (6) and (7). Taking bases of the iterated kernels $E_{\ell_j}(\lambda_j)$ and placing them as columns of a matrix P we reach a block diagonal form.

Let us show that the dimensions of the iterated kernels of A_1 coincide with the dimensions of $E_j(\lambda_1)$. Let $B = P^{-1}\,A\,P$. Since

$$u \in \ker(B - \lambda_1 I)^j \qquad \Longleftrightarrow \qquad P\,u \in E_j(\lambda_1) \subseteq E_{\ell_1}(\lambda_1) = \operatorname{span}[p_1, p_2, \dots, p_{m_1}]$$

(p_j are the column vectors of P), then

$$\ker(B - \lambda_1 I)^j \subseteq \operatorname{span}[e_1, e_2, \dots, e_{m_1}]$$

(e_j are the canonical vectors). This means that elements of $\ker(B - \lambda_1 I)^j$ can only have its first m_1 entries non-vanishing. From that, it is obvious that

$$u \in \ker(B - \lambda_1 I)^j \qquad \Longleftrightarrow \qquad u = \begin{bmatrix} v \\ 0 \end{bmatrix}, \quad v \in \ker(A_1 - \lambda_1 I)^j.$$

The same proof applies to all other eigenvalues. □

Very relevant conclusion. Once the first Jordan Theorem has been proved, we can restrict our attention to matrices whose characterstic polynomial is of the form $(\lambda_1 - x)^n$, since this theorem allows us to separate the space into smaller subspaces where the operator/matrix only involves one eigenvalue.

3 Proofs

This section can be skipped in a first reading.

Property (5). For all $j \geq 1$

$$\dim E_{j+1}(\lambda) - \dim E_j(\lambda) \leq \dim E_j(\lambda) - \dim E_{j-1}(\lambda).$$

Proof. Let $m = \dim E_{j+1}(\lambda) - \dim E_j(\lambda)$ and let us take independent vectors $\{v_1, \ldots, v_m\}$ such that

$$E_{j+1}(\lambda) = E_j(\lambda) \oplus \operatorname{span}[v_1, \ldots, v_m].$$

This can be written in the following form:

$$\begin{array}{lcl} v_1, \ldots, v_m \in E_{j+1}(\lambda), & & \\ \xi_1 v_1 + \ldots + \xi_m v_m \in E_j(\lambda) & \iff & \xi_1 = \ldots = \xi_m = 0. \end{array}$$

Let now $w_i := (A - \lambda I)v_i$. It is easy to verify that

$$\begin{array}{lcl} w_1, \ldots, w_m \in E_j(\lambda), & & \\ \xi_1 w_1 + \ldots + \xi_m w_m \in E_{j-1}(\lambda) & \iff & \xi_1 = \ldots = \xi_m = 0, \end{array}$$

and therefore

$$\begin{array}{ll} \dim \operatorname{span}[w_1, \ldots, w_m] = m & \text{(the vectors are independent)}, \\ \operatorname{span}[w_1, \ldots, w_m] \subset E_j(\lambda), & \\ \operatorname{span}[w_1, \ldots, w_m] \cap E_{j-1}(\lambda) = \{0\}. & \end{array}$$

Therefore $\dim E_j(\lambda) - \dim E_{j-1}(\lambda) \geq m$. □

Lemma 1. *If $p(A)u = 0$ and $q(A)u = 0$, where $p, q \in \mathbb{C}[x]$, and we define $r = \text{g.c.d.}(p, q)$, then $r(A)u = 0$.*

Proof. Both p and q have to be multiples of the minimal polynomial of u with respect to A. Therefore r is a multiple of this same polynomial which implies the result. (Recall that the minimal polynomial for a vector u with respect to a matrix A is the lowest order monic polynomial p such that $p(A)u = 0$ and that if $q(A)u = 0$, then p divides q.) □

Property (6). Let $\lambda_1, \ldots, \lambda_k$be pairwise different eigenvalues of A. Let $j_1, \ldots, j_k \geq 1$. The the following sum of spaces is direct:

$$E_{j_1}(\lambda_1) \oplus E_{j_2}(\lambda_2) \oplus \ldots \oplus E_{j_k}(\lambda_k).$$

Proof. Let $u_1, \ldots, u_k$ such that

$$u_1 + u_2 + \ldots + u_k = 0, \qquad u_i \in E_{j_i}(\lambda_i).$$

Then

$$\begin{array}{l} (A - \lambda_1 I)^{j_1} u_1 = 0, \\ (A - \lambda_2 I)^{j_2} \ldots (A - \lambda_k I)^{j_k} u_1 = -(A - \lambda_2 I)^{j_2} \ldots (A - \lambda_k I)^{j_k} (u_2 + \ldots + u_k) = 0. \end{array}$$

Since

$$\text{g.c.d.}\Big((x-\lambda_1)^{j_1},\ (x-\lambda_2)^{j_2} \dots (x-\lambda_k)^{j_k} \Big) = 1,$$

then $u_1 = 0$. Proceeding by induction we prove that $u_j = 0$ for all j. This proves that the subspaces are independent. □

Property (7a). For all j

$$\dim E_j(\lambda) \le m_\lambda,$$

where $\chi_A(x) = (x-\lambda)^{m_\lambda} q(x)$ with $q(\lambda) \neq 0$.

Proof. Let P be invertible such that its first m columns form a basis of $E_j(\lambda)$ ($m = \dim E_j(\lambda)$). Therefore

$$P^{-1}\,A\,P = \begin{bmatrix} A_m & C_m \\ 0 & B_m \end{bmatrix}, \qquad A_m \in \mathcal{M}(m;\mathbb{C}).$$

If $A_m\,v = \mu\,v$, then

$$P^{-1}\,A\,P \begin{bmatrix} v \\ 0 \end{bmatrix} = \mu \begin{bmatrix} v \\ 0 \end{bmatrix} \qquad \Longrightarrow \qquad A\,P \begin{bmatrix} v \\ 0 \end{bmatrix} = \mu P \begin{bmatrix} v \\ 0 \end{bmatrix}.$$

Therefore

$$w = P \begin{bmatrix} v \\ 0 \end{bmatrix} \in E_1(\mu) \cap E_j(\lambda).$$

If $\mu \neq \lambda$, then $E_1(\mu) \cap E_j(\lambda) = \{0\}$, so $w = 0$ and $v = 0$. This shows that $\chi_{A_m}(x) = (\lambda - x)^m$. Since

$$\chi_A(x) = \chi_{A_m}(x)\chi_{B_m}(x) = (\lambda-x)^m \chi_{B_m}(x) = (\lambda-x)^{m_\lambda} q(x), \qquad q(\lambda) \neq 0$$

it follows that $m \le m_\lambda$. □

Property (7b). If $E_{\ell+1}(\lambda) = E_\ell(\lambda)$, then

$$\dim E_\ell(\lambda) \ge m_\lambda.$$

Proof. Let $m = \dim E_\ell(\lambda)$. We take P as in the previous proof, so that

$$P^{-1}\,A\,P = \begin{bmatrix} A_m & C_m \\ 0 & B_m \end{bmatrix} = \widetilde{A},$$

where $A_m \in \mathcal{M}(m;\mathbb{C})$. We are going to see that B_m cannot have λ as an eigenvalue.

Let us recall that for all j

$$u \in \ker(\widetilde{A} - \lambda\,I)^j \qquad \Longleftrightarrow \qquad P\,u \in \ker(A - \lambda\,I)^j.$$

Therefore

$$\ker(\widetilde{A} - \lambda\,I)^\ell = \ker(\widetilde{A} - \lambda\,I)^{\ell+k}, \qquad \forall k$$
$$\ker(\widetilde{A} - \lambda\,I)^\ell = \operatorname{span}[e_1, \dots, e_m].$$

The latter identity implies that the first m columns of $(\widetilde{A}-\lambda\, I)^\ell$ vanish. We can then write

$$(\widetilde{A}-\lambda\, I)^\ell = \begin{bmatrix} (A_m-\lambda\, I)^\ell & \widehat{C}_m \\ 0 & (B_m-\lambda\, I)^\ell \end{bmatrix} = \begin{bmatrix} 0 & \widehat{C}_m \\ 0 & (B_m-\lambda\, I)^\ell \end{bmatrix}.$$

Now, if $(B_m-\lambda I)v=0$, then

$$(\widetilde{A}-\lambda\, I)^{2\ell}\begin{bmatrix} 0 \\ v \end{bmatrix} = \begin{bmatrix} 0 & \widehat{C}_m \\ 0 & (B_m-\lambda\, I)^\ell \end{bmatrix}\begin{bmatrix} \widehat{C}_m v \\ (B_m-\lambda\, I)^\ell v \end{bmatrix} = \begin{bmatrix} 0 & \widehat{C}_m \\ 0 & (B_m-\lambda\, I)^\ell \end{bmatrix}\begin{bmatrix} \widehat{C}_m v \\ 0 \end{bmatrix} = 0$$

which implies

$$(\widetilde{A}-\lambda\, I)^\ell \begin{bmatrix} 0 \\ v \end{bmatrix} = 0$$

and therefore $v=0$. □

Nota. Properties (7a) and (7b) imply Property (7). However, Property (7b) is enough to show that $\dim E_\ell(\lambda)=m(\lambda)$. The reason for this is that Property (6) implies

$$n \ge \dim E_{\ell_1}(\lambda_1)+\ldots+\dim E_{\ell_k}(\lambda_k) \ge m_1+\ldots+m_k = n$$

which forces all the inequalities to be equalities.

4 Jordan blocks and Jordan matrices

Definition 4. *A Jordan block of order k associated to a value λ is a $k\times k$ matrix of the form*

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}.$$

Properties.

(1) The characteristic polynomial of $J_k(\lambda)$ is $(\lambda-x)^k$. Moreover,

$$\dim\ker(J_k(\lambda)-\lambda I)=1.$$

(2) The dimensions of the iterated kernels for $J_k(\lambda)$ are:

$$\dim\ker(J_k(\lambda)-\lambda I)^j = \left\{ \begin{array}{ll} j, & j=1,\ldots,k \\ k, & j\ge k \end{array} \right\} = \min(k,j).$$

Proof. Note that

$$J_k(\lambda) - \lambda I = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$$

and take increasing powers of this matrix, noticing how the rank reduces by onw with each power. Otherwise, note that $n_1 = 1$ (the dimension of the space of eigenvectors is one) and by Property (5) of the iterated kernels the dimensions can only increase one by one until it reaches k. □

Definition 5. *A Jordan matrix associated to a value λ is a block diagonal matrix whose blocks are Jordan blocks for the value λ:*

$$\begin{bmatrix} J_{k_1}(\lambda) & & & \\ & J_{k_2}(\lambda) & & \\ & & \ddots & \\ & & & J_{k_r}(\lambda) \end{bmatrix}$$

A general Jordan matrix is a block diagonal matrix, whose blocks are Jordan matrices associated to different values. Equivalently, it is a block diagonal matrix whose blocks are Jordan blocks.

5 Two combinatorial problems

Objetivo. We will next show how to establish a one-to-one correspondence between the different possibilities for dimensions of iterated kernels for an $m \times m$ Jordan matrix (with a single eigenvalue) and the possibilities of setting up the Jordan boxes until we fill the $m \times m$ space.

Problem (P). Find a sequence of integers satisfying

(a) $0 = n_0 < n_1 < \ldots < n_\ell = n_{\ell+1} = \ldots,$

(b) $n_\ell - n_{\ell-1} \leq n_{\ell-1} - n_{\ell-2} \leq \ldots \leq n_2 - n_1 \leq n_1 - n_0 = n_1$

(c) $n_\ell = m.$

Problem (Q). Find a sequence of non-negative integers $q_1, q_2, \ldots, q_\ell, \ldots$ such that

(a) $q_{\ell+1} = q_{\ell+2} = \ldots = 0$

(b) $q_1 + 2\,q_2 + 3\,q_3 + \ldots + \ell\,q_\ell = m$

Remark. Before showing how Problem (P) and Problem (Q) are related, let us look ahead and explain what they are going to mean in the context of Jordan matrices. We also give the interpretation for the solutions of an intermediate problem:

- $\{n_j\}$ will be the dimensions of the iterated kernels;
- $\{p_j\}$ will be the number of Jordan blocks of size $j \times j$ or higher;
- $\{q_j\}$ will be the number of $j \times j$ Jordan blocks.

The relation between both sequences will be done without taking into account what is the value of ℓ where the solution of (P) stagnates and where the one of (Q) takes its last non-zero value. This value will be however the same in both sequences. □

From (P) to (Q). Given a solution to (P), we define

$$p_j = n_j - n_{j-1}, \qquad j \geq 1.$$

This sequence satisfies

$$0 = \ldots = p_{\ell+1} < p_\ell \leq p_{\ell+1} \leq \ldots \leq p_1 = n_1.$$

We next define

$$q_j = p_j - p_{j+1}, \qquad j \geq 1,$$

so that $q_\ell = p_\ell$ and $q_j = 0$ for all $j > \ell$. The sequence $\{q_j\}$ is a solution to (Q).

From (Q) to (P). Given a solution to (Q), we define

$$\begin{array}{ll} p_j = q_j + q_{j+1} + \ldots + q_\ell, & \qquad j = 1, \ldots, \ell, \\ p_j = 0 & \qquad j \geq \ell, \end{array}$$

and then define a new sequence using a recurrence

$$\begin{cases} n_0 = 0 \\ n_j = p_j + n_{j-1}, & \qquad j \geq 1. \end{cases}$$

Then $\{n_j\}$ is a solution of (P). This shows that there is a bijection between solutions of (P) and (Q).

Theorem 2. *Let $\{n_j\}$ be a solution to* (P) *and let $\{q_j\}$ be the associated solution to* (Q)*. Then the $m \times m$ Jordan matrix associated to the value λ, built with*

$$q_j \textit{ blocks of size } j \times j \textit{ for all } j$$

satisfies

$$\dim E_j(\lambda) = n_j, \qquad j \geq 0.$$

Proof. Let us consider the Jordan matrix constructed as explained in the statement. Since the matrix is block diagonal, it follows that

$$\begin{aligned}
\dim E_j(\lambda) &= q_\ell \dim\ker(J_\ell(\lambda)-\lambda I)^j + q_{\ell-1} \dim\ker(J_{\ell-1}(\lambda)-\lambda I)^j \\
&\quad + \ldots + q_1 \dim\ker(J_1(\lambda)-\lambda I)^j \\
&= q_\ell \min(j,\ell) + q_{\ell-1}\min(j,\ell-1) + \ldots + q_1 \min(j,1) \\
&= j\,(q_\ell + q_{\ell-1} + \ldots + q_{j+1}) + j q_j + (j-1) q_{j-1} + \ldots + 2q_2 + q_1 \\
&= j(p_\ell + p_{\ell-1} - p_\ell + \ldots + p_{j+1} - p_{j+2}) \\
&\quad + j(p_j - p_{j+1}) + \ldots + 2(p_2 - p_3) + (p_1 - p_2) \\
&= j\,p_{j+1} - j\,p_{j+1} + p_j + p_{j-1} + \ldots + p_1 \\
&= n_j.
\end{aligned}$$

This proves the result. □

Conclusions.

(1) For any solution of problem (P), there exists an $m \times m$ Jordan matrix with a single eigenvalue such that

$$\dim E_j(\lambda) = n_j, \qquad \forall j.$$

(2) Consequently, given possible configurations of the dimensions of the iterated kernels associated to all eigenvalues, there exists a matrix (a Jordan matrix actually) whose iterated kernels have the given dimensions.

(3) Two essentially different Jordan matrices (that cannot be obtained by permutation in the order of the blocks) are not similar.

Proof. The block-configuration (how many blocks of each size for each eigenvalue) determines univocally the dimensions of the iterated kernels (different block configurations imply different dimensions of the kernels). However, the dimensions of the iterated kernel do not vary under similarity transformations. □

6 Existence of a Jordan canonical form

An observation. If

$$P^{-1} A\, P = \begin{bmatrix} J_k(\lambda) & \times \\ 0 & \times \end{bmatrix}$$

and $u_1, \ldots, u_k$ are the first columns of P, then

$$\begin{aligned}
Au_1 &= \lambda u_1 \\
Au_2 &= \lambda u_2 + u_1 \\
&\vdots \\
Au_k &= \lambda u_k + u_{k-1}
\end{aligned}$$

so for all $j \geq 2$

$$u_{j-1} = (A - \lambda I) u_j.$$

Lemma 2. *If $\mathcal{V}$ is a subspace satisfying*

$$E_{j-1}(\lambda) \oplus \mathcal{V} \subseteq E_j(\lambda)$$

and $\mathcal{W} = (A - \lambda I)\mathcal{V} = \{(A - \lambda I)\, u \, : \, u \in \mathcal{V}\}$, *then*

$$E_{j-2}(\lambda) \oplus \mathcal{W} \subseteq E_{j-1}(\lambda), \qquad \dim \mathcal{W} = \dim \mathcal{V}.$$

Proof. There are three things to check: (a) $\mathcal{W} \subset E_{j-1}(\lambda)$, which is quite obvious, since

$$u \in E_j(\lambda) \qquad \Longrightarrow \qquad (A - \lambda I)u \in E_{j-1}(\lambda);$$

(b) $\mathcal{W} \cap E_{j-2}(\lambda) = 0$, since if $v \in \mathcal{W}$ (so $v = (A - \lambda I)u$, for some $u \in \mathcal{V}$), then

$$(A - \lambda I)^{j-2} v = 0 \qquad \Longrightarrow \qquad (A - \lambda I)^{j-1} u = 0 \qquad \Longrightarrow \qquad u \in \mathcal{V} \cap E_{j-1}(\lambda) = 0;$$

(c) if $u \in \mathcal{V}$, then

$$(A - \lambda I)u = 0 \qquad \Longrightarrow \qquad u \in E_1(\lambda) \cap \mathcal{V} \subseteq E_{j-1}(\lambda) \cap \mathcal{V} = 0,$$

so linear independence is preserved by multiplication by $(A - \lambda I)$. □

The proof of Theorem 3 is quite technical. It can be skipped in a first reading.

Theorem 3 (Second Jordan decomposition theorem)**.** *Every $m \times m$ with characteristic polynomial $(x - \lambda)^m$ is similar to a Jordan matrix associated to the value λ. Moreover, the dimensions of the Jordan boxes are given by the coefficients $\{q_j\}$ associated to the dimensions $\{n_j\}$ of the iterated kernels.*

Proof. Let $E_\ell(\lambda) = \mathbb{C}^m$ be the first iterated kernel of maximum dimension. Let $\mathcal{V}_\ell = \mathcal{V}_\ell^0$ tal que

$$E_\ell(\lambda) = E_{\ell-1}(\lambda) \oplus \mathcal{V}_\ell^0,$$

so that $\dim \mathcal{V}_\ell^0 = n_\ell - n_{\ell-1} = p_\ell = q_\ell$. Using Lemma 2 above

$$E_{\ell-1}(\lambda) = E_{\ell-2}(\lambda) \oplus \underbrace{(A - \lambda I)\mathcal{V}_\ell^0 \oplus \mathcal{V}_{\ell-1}^0}_{\mathcal{V}_{\ell-1}},$$

with

$$\dim \mathcal{V}_{\ell-1}^0 = (n_{\ell-1} - n_{\ell-2}) - \dim \mathcal{V}_\ell^0 = p_{\ell-1} - p_\ell = q_{\ell-1} \geq 0.$$

Likewise

$$\begin{aligned} E_{\ell-2}(\lambda) &= E_{\ell-3}(\lambda) \oplus \underbrace{(A - \lambda I)\mathcal{V}_{\ell-1} \oplus \mathcal{V}_{\ell-2}^0}_{\mathcal{V}_{\ell-2}} = \\ &= E_{\ell-3}(\lambda) \oplus (A - \lambda I)^2 \mathcal{V}_\ell^0 \oplus (A - \lambda I)\mathcal{V}_{\ell-1}^0 \oplus \mathcal{V}_{\ell-2}^0, \end{aligned}$$

with

$$\dim \mathcal{V}^0_{\ell-2} = q_{\ell-2}.$$

Proceeding succesively, we can build a partition of $E_\ell(\lambda)$

$$\begin{array}{ccccccccccc}
E_\ell(\lambda) & = & \mathcal{V}^0_\ell & \oplus & (A-\lambda I)\mathcal{V}^0_\ell & \oplus & \dots & \oplus & (A-\lambda I)^{\ell-2}\mathcal{V}^0_\ell & \oplus & (A-\lambda I)^{\ell-1}\mathcal{V}^0_\ell \\
 & & & \oplus & \mathcal{V}^0_{\ell-1} & \oplus & \dots & \oplus & (A-\lambda I)^{\ell-3}\mathcal{V}^0_{\ell-1} & \oplus & (A-\lambda I)^{\ell-2}\mathcal{V}^0_{\ell-1} \\
 & & & & & & & \vdots & & \vdots & \\
 & & & & & & & \oplus & \mathcal{V}^0_2 & \oplus & (A-\lambda I)\mathcal{V}^0_2 \\
 & & & & & & & & & \oplus & \mathcal{V}^0_1
\end{array}$$

so that,

$$\dim \mathcal{V}^0_j = q_j, \qquad j = 1, \dots, \ell.$$

Note that we can decompose $E_\ell()\lambda)$ in two different ways:

$$E_\ell(\lambda) = \mathcal{V}_\ell \oplus \mathcal{V}_{\ell-1} \oplus \dots \oplus \mathcal{V}_2 \oplus \mathcal{V}_1 = \mathcal{W}_\ell \oplus \mathcal{W}_{\ell-1} \oplus \dots \oplus \mathcal{W}_2 \oplus \mathcal{W}_1,$$

where

$$\mathcal{V}_j = \mathcal{V}^0_j \oplus (A-\lambda I)\mathcal{V}^0_{j+1} \oplus \dots \oplus (A-\lambda I)^{\ell-j}\mathcal{V}^0_\ell \qquad \text{(add by columns)}$$

and

$$\mathcal{W}_j = \mathcal{V}^0_j \oplus (A-\lambda I)\mathcal{V}^0_j \oplus \dots \oplus (A-\lambda I)^{j-1}\mathcal{V}^0_j \qquad \text{(add by rows)}$$

We next build a basis for $E_\ell(\lambda)$ in the following form. If $\mathcal{V}^0_j \neq \{0\}$, we start

$$\{u_{j,1}, \dots, u_{j,q_j}\} \qquad \text{basis for } \mathcal{V}^0_j.$$

We then create sequences of j vectors for each of the above vectors. If u is any of them, we define

$$\begin{array}{rll}
v_1 = (A-\lambda I)^{j-1}u = (A-\lambda I)v_2 & & \in (A-\lambda I)^{j-1}\mathcal{V}^0_j \\
v_2 = (A-\lambda I)^{j-2}u = (A-\lambda I)v_3 & & \in (A-\lambda I)^{j-2}\mathcal{V}^0_j \\
\vdots & & \\
v_{j-1} = (A-\lambda I)u \quad = (A-\lambda I)v_j & & \in (A-\lambda I)\mathcal{V}^0_j \\
v_j = u, & & \in \mathcal{V}^0_j,
\end{array}$$

so that

$$Av_1 = \lambda v_1, \qquad Av_k = v_{k-1} + \lambda v_k, \quad k = 2, \dots, j.$$

We thus obtain a basis for

$$\mathcal{W}_j = (A-\lambda I)^{j-1}\mathcal{V}^0_j \oplus \dots \oplus (A-\lambda I)\mathcal{V}^0_j \oplus \mathcal{V}^0_j.$$

If we change basis to this new basis (collect the corresponding basis for each row of the large decomposition above, noting that some rows might be zero), we have a Jordan matrix with q_j blocks of size $j \times j$ for each j. □

Theorem 4. *Every matrix is similar to a Jordan matrix (its Jordan canonical form). The Jordan form is unique up to permutation of its blocks, and it is the only general Jordan matrix such that the dimensions of the iterated kernels for all the eigenvalues coincide with those of A.*

Proof. Use the First Jordan Theorem to separate eigenvalues, and then use the construction of the second theorem. □

Corollary 1. *The similarity class of a matrix is determined by the characteristic polynomials and the dimensions of all the iterated kernels (for all roots of the characteristic polynomial). Matrices with different Jordan forms are not similar.*

7 Additional topics

7.1 Cayley–Hamilton's Theorem

Theorem 5. *For every matrix*

$$\chi_A(A) = 0.$$

Proof. A simple form of proving this theorem is using the Jordan form. Note that there are other proofs using much less complicated results. Let J be the Jordan form for A. Since $\chi_J = \chi_A$, we only need to prove that $\chi_J(J) = 0$. Hoever

$$\chi_J(J) = \begin{bmatrix} \chi_J(J_{k_1}(\lambda_1)) & & & \\ & \chi_J(J_{k_2}(\lambda_2)) & & \\ & & \ddots & \\ & & & \chi_J(J_{k_r}(\lambda_r)) \end{bmatrix}$$

and on the other hand

$$(J_k(\lambda) - \lambda I)^k = 0,$$

which proves the theorem. □

Another proof. It is possible to prove this same result using the First Jordan decomposition theorem. Using this theorem and the fact that

$$p(P^{-1}AP) = \begin{bmatrix} p(A_1) & & & \\ & p(A_2) & & \\ & & \ddots & \\ & & & p(A_k) \end{bmatrix},$$

it is enough to prove the result for matrices with only one eigenvalue. Now, if the characteristic polynomial of A is $(x-\lambda)^n$, there exists $\ell \le n$ such that $E_\ell(\lambda) = \ker, (A - \lambda I)^\ell = \mathbb{C}^n$, which implies that $(A - \lambda I)^\ell = 0$ with $\ell \le n$. □

Some comments. The minimal polynomial for A is the lowest degree monic polynomial satisfying $Q(A) = 0$. Assume that

$$\chi_A(x) = (\lambda_1 - x)^{m_1} \dots (\lambda_k - x)^{m_k}.$$

Since the minimal polynomial for A and its Jordan form J are the same, it is easy to see that if for every λ_j we pick the minimum ℓ_j such that

$$\dim E_{\ell_j}(\lambda_j) = m_j,$$

then the minimal polynomial is

$$(x - \lambda_1)^{\ell_1} \dots (x - \lambda_k)^{\ell_k}.$$

Another simple argument shows then that the following statements are equivalent:

(a) the matrix is diagonalizable

(b) the Jordan form of the matrix is diagonal (all blocks are 1×1)

(c) all iterated kernels have maximum dimension

$$E_2(\lambda) = E_1(\lambda), \qquad \forall \lambda \text{ eigenvalue of } A.$$

(d) the minimal polynomial has simple roots.

7.2 The transposed Jordan form

In many contexts it is common to meet a Jordan form made up of blocks of the form

$$\begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{bmatrix},$$

that is, with the 1s under the main diagonal. If we have been able to get to the usual Jordan form, we only need to reoder the basis to change to this transposed form, namely, if the vectors

$$u_1, u_2, \dots, u_p$$

correspond to a $p \times p$ block associated to λ, then the vectors

$$u_p, u_{p-1}, \dots, u_1$$

give the same block with the one elements under the main diagonal.

Corollary 2. *Every matrix is similar to its transpose.*

Proof. A simple reordering of the basis shows that every Jordan matrix is similar to its transpose. Also, the tranpose of a matrix is similar to the transpose of its Jordan form. This finishes the proof. □

7.3 The real Jordan form

If $A \in \mathcal{M}(n;\mathbb{R})$, $\lambda = \alpha + \beta\imath$ (with $\beta \neq 0$) is an eigenvalue and we have the basis for $E_\ell(\lambda)$ giving the Jordan blocks for λ, then we just need to conjugate the vectors of the basis to obtain the basis giving the Jordan blocks for $\overline{\lambda} = \alpha - \beta\imath$. The reason is simple:

$$u \in E_j(\lambda) \qquad \Longleftrightarrow \qquad \overline{u} \in E_j(\overline{\lambda})$$

This means that the configuration of Jordan blocks for λ and $\overline{\lambda}$ are the same.

If $u_1, \ldots, u_p$ are the basis vectors associated to $\lambda = \alpha + \beta\imath$ (with $\beta \neq 0$), for a basis leading to Jordan form, then $\overline{u_1}, \ldots, \overline{u_p}$ can be taken as the vectors associated to $\overline{\lambda}$. We can then take

$$\begin{aligned} v_1 &= \operatorname{Re} u_1, & v_2 &= \operatorname{Im} u_1 \\ v_3 &= \operatorname{Re} u_2, & v_4 &= \operatorname{Im} u_2 \\ &\vdots \\ v_{2p-1} &= \operatorname{Re} u_p, & v_{2p} &= \operatorname{Im} u_p \end{aligned}$$

(Re and Im denote the real and imaginary part of the vector). With these vectors, the blocks corresponding to λ and $\overline{\lambda}$ mix up, producing bigger blocks with real numbers. For instance

$$\left.\begin{matrix} \begin{bmatrix} \alpha+\beta\imath & 1 \\ 0 & \alpha+\beta\imath \end{bmatrix} \\ \\ \begin{bmatrix} \alpha-\beta\imath & 1 \\ 0 & \alpha-\beta\imath \end{bmatrix} \end{matrix}\right\} \longmapsto \begin{bmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix}$$

In this way, when a real matrix has complex eigenvalues, we can restrict our bases to have real components and work with block-based Jordan blocks

$$J_k(\lambda, \overline{\lambda}) := \begin{bmatrix} \Lambda & I_2 & & \\ & \Lambda & \ddots & \\ & & \ddots & I_2 \\ & & & \Lambda \end{bmatrix}, \qquad \Lambda := \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \qquad I_2 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

7.4 Symmetric matrices

Theorem 6. *Every real symmetric matrix is diagonalizable.*

Proof. Asssume that $A \in \mathcal{M}(n;\mathbb{R})$ satisfyes $A^{Tr} = A$. Let $Au = \lambda u$, with $u \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$. Then

$$\lambda\, \|u\|^2 = \overline{u}^{Tr} A\, u = \overline{(A\, u)}^{Tr} u = \overline{\lambda} \|u\|^2,$$

which proves that λ is real. We now only need to prove that $E_2(\lambda) = E_1(\lambda)$. This is a consequence of the following chain of implications:

$$\begin{array}{lll} u \in E_2(\lambda) & \Longrightarrow & (A-\lambda I)^2 u = 0 \\ & \Longrightarrow & u^{Tr}(A-\lambda I)^2 u = 0 \\ & \Longrightarrow & ((A-\lambda I)u)^{Tr}((A-\lambda I)u) = 0 \\ & \Longrightarrow & \|(A-\lambda I)u\|^2 = 0 \\ & \Longrightarrow & u \in E_1(\lambda). \end{array}$$

□

8 Exercises

1. (Section 1) Show that if $BA = AB$, then $\ker B$ and $\operatorname{range} B$ are $A-$invariant.

2. (Section 1) Let $\mathcal{V}$ be $A-$invariant and let $\mathcal{W}$ be such that $\mathbb{F}^n = \mathcal{V} \oplus \mathcal{W}$. Show that A is similar to a block upper triangular matrix. (Hint. Take a basis of $\mathbb{F}^n$ composed of vectors of $\mathcal{V}$ followed by vectors of $\mathcal{W}$.)

3. (Section 1) Let A be such that

$$P^{-1}AP = \begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}$$

and let $\{p_1, \ldots, p_4\}$ be the columns of P. Show that $\operatorname{span}[p_1], \operatorname{span}[p_1, p_2]$ and $\operatorname{span}[p_1, p_2, p_3]$ are $A-$invariant.

4. (Section 2) We can define the iterated ranges as follows

$$R_j(\lambda) = \operatorname{range}(A - \lambda I)^j, \qquad j \geq 1$$

with $R_0(\lambda) = \operatorname{range}(I) = \mathbb{C}^n$. Show that

   (a) $R_{j+1}(\lambda) \subseteq R_j(\lambda)$ for all j.

   (b) If $R_{\ell+1}(\lambda) = R_\ell(\lambda)$, then $R_{\ell+2}(\lambda) = R_{\ell+1}(\lambda)$. (Hint. This result can be proved directly or using the equivalent result for the iterated kernels.)

   (c) $\dim R_{j+1}(\lambda) - \dim R_j(\lambda) \geq \dim R_j(\lambda) - \dim R_{j-1}(\lambda)$.

If $R_{\ell+1}(\lambda) = R_\ell(\lambda)$, what is the dimension of this subspace?

5. (Section 4) Let A be such that

$$P^{-1}AP = J_n(\lambda)$$

and let $\{p_1, p_2, \ldots, p_n\}$ be the column vectors of P.

(a) Show that $p_j \in E_j(\lambda)$.

(b) Show that $(A - \lambda I)^{j-1} p_j = p_1$.

(c) Use the previous results to find the minimal polynomial of p_j.

(d) Use (c) to find the minimal polynomial of A.

6. (Sections 5 and 6) In a 12-dimensional space, we let

$$q_1 = 2, \quad q_2 = 0, \quad q_3 = 2, \quad q_4 = 1, \qquad q_j = 0 \quad \forall j \geq 5.$$

(This corresponds to a Jordan matrix with two 1×1 blocks, two 3×3 blocks and one 4×4 blocks.) What is the associated solution of Problem (P)? In other words, what are the dimensions of the iterated kernels $E_j(\lambda)$ for a matrix with the Jordan structure given above?

7. (Section 6) Let A be such that

$$\chi_A(x) = (2 - x)^3 (4 - x)^2.$$

List all possible Jordan forms for A. List the associated minimal polynomial. (Hint. It is easy to find the minimal polynomial for all the vectors in the basis that leads to the Jordan form. The l.c.m. of the minimal polynomials for the vectors in any basis is the minimal polynomial for the matrix.)

8. (Section 6) Let A be such that

$$\chi_A(x) = (3 - x)^4 (2 - x)^2 \qquad \text{and} \qquad \mathrm{minP}_A(x) = (x - 3)^2 (x - 2)^2.$$

How many possible Jordan forms can A have?

9. (Section 6) Let A be such that

$$\chi_A(x) = (2 - x)^6.$$

Write down all possible Jordan forms for A. For each of them, list the dimensions of the iterated kernels and the minimal polynomial of the matrix.

10. (Section 7) Prove that every complex Hermitian matrix has only real eigenvalues and is diagonalizable. (Hint. Instead of transposing, use conjugate transposition in the proof of Theorem 6.)

References

[1] Hernández, Eugenio. *Algebra and Geometry* (Spanish), Addison-Wesley.

[2] Kress, Rainer. *Linear Integral Operators*. Springer Verlag.

[3] Katznelson, Yitzhak, and Katznelson, Yonatan R. *A (Terse) Introduction to Linear Algebra*. AMS Student Mathematical Library.

---

by F.–J. Sayas

Mathematics, Ho!, October 2002 – Revised and translated November 2013