The Real Jordan Canonical Form and the Rational Canonical Form

Not all matrices over a given field have a Jordan canonical form as not all polynomials split completely into linear factors. For example, over the reals one can have irreducible quadratic factors. A field in which every polynomial splits completely into a product of linear factors over that field is said to be algebraically closed. An example, of such a field is the field of complex numbers.

Let A be an $n \times n$ real matrix and let $T = T_A : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ be the associated linear operator. The minimal polynomial $m_A(X)$ of A has the form

$$(X - c_1)^{m_1} \cdots (X - c_k)^{m_k}((X - a_1)^2 + b_1^2)^{n_1} \cdots ((X - a_\ell)^2 + b_\ell^2)^{n_\ell}$$

which has the factorization

$$(X - c_1)^{m_1} \cdots (X - c_k)^{m_k}(X - \alpha_1)^{n_1}(X - \overline{\alpha}_1)^{n_1} \cdots (X - \alpha_\ell)^{n_\ell}(X - \overline{\alpha}_\ell)^{n_\ell}$$

over $\mathbb{C}$ with $\alpha_j = a_j + b_j i$, $\overline{\alpha}_j = a_j - b_j i$. Since the nullity of a real matrix is the same as the nullity of the matrix viewed as a complex matrix, the number of Jordan blocks $J_r(\lambda)$ with λ real can be computed as before. The complex Jordan blocks come in conjugate pairs $J = J_r(a+ib)$, $\overline{J} = J_r(a-ib)$ and their number of such pairs is $s_r/2$ where

$$s = (s_0, s_1, \ldots, s_j, \ldots) = -R(L-1)^2(d),$$

where $d = (d_0, d_1, \ldots, d_j, \ldots)$ with $d_j = \mathrm{nullity}((\mathrm{A}-\mathrm{a})^2 + \mathrm{b}^2)^{\mathrm{j}}$. If $u \in \mathbb{C}^{n\times 1}$ is the cyclic vector associated to $J_r(a + bi)$ then $\overline{u}$ is a cyclic vector for $J_r(a - ib)$. If

$$g = \frac{u - \overline{u}}{2i}, \quad h = \frac{u + \overline{u}}{2},$$

then g, h are real vectors such that, over $\mathbb{C}$,

$$W = \mathrm{Span}(\mathrm{g}, \mathrm{h}, \mathrm{T}(\mathrm{g}), \mathrm{T}(\mathrm{h}), \ldots, \mathrm{T}^{\mathrm{r}-1}(\mathrm{g}), \mathrm{T}^{\mathrm{r}-1}(\mathrm{h})) = \mathrm{Span}(\mathrm{u}, \overline{\mathrm{u}}, \mathrm{T}(\mathrm{u}), \mathrm{T}(\overline{\mathrm{u}}), \ldots, \mathrm{T}^{\mathrm{r}-1}(\mathrm{u}), \mathrm{T}^{\mathrm{r}-1}(\overline{\mathrm{u}}))$$

If $u_1 = (T - \alpha)^{r-1}(u), \ldots, u_{r-1} = (T - \alpha)(u), u_r = u$ and $\alpha = a + bi$ then

$$T(u_1) = \alpha u_1, \quad T(u_j) = \alpha u_j + u_{j-1} \text{ for } 2 \leq j \leq r.$$

Setting $f_{2j-1} = \frac{u_j - \overline{u}_j}{2i}$, $f_{2j} = \frac{u_j + \overline{u}_j}{2}$ for $1 \leq j \leq r$, we get

$$T(f_{2j-1}) = af_{2j-1} + bf_{2j} + f_{2j-3}, \quad T(f_{2j}) = -bf_{2j-1} + bf_{2j} + f_{2j-2}$$

for $2 \leq j \leq r$ and $T(f_1) = af_1 + bf_2$, $T(f_2) = -bf_1 + af_2$. The real vectors $f_1, \ldots, f_{2r}$ are linearly independent and span W. The matrix of the restriction of T to the subspace of $\mathbb{R}^{n\times 1}$ spanned by $f_1, \ldots, f_{2r}$ with respect to $f = (f_1, \ldots, f_{2r})$ is the $2r \times 2r$ block matrix $RJ_r(a, b) = [B_{ij}]$ with

$$B_{ii} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad B_{i(i+1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and B_{ij} the zero 2×2 matrix for all other i, j. We thus obtain the following theorem:

Theorem. If T is a linear operator on a finite-dimensional real vector space then there is a basis for V such that the matrix of T with respect to this basis has a block diagonal form with blocks of the form $J_k(a)$ or $RJ_k(a,b)$. Moreover, if s_k is the number of blocks of the form $J_k(a)$, then

$$s = (s_0, s_1, \ldots, s_i, \ldots) = -R(L-1)^2(d)$$

where $d = (d_0, d_1, \ldots, d_i, \ldots)$ with $d_i = \dim \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}})$. If s_k is the number of blocks of the form $RJ_k(a,b)$, then

$$s = (s_0, s_1, \ldots, s_i, \ldots) = -R(L-1)^2(d)$$

where $d = (d_0, d_1, \ldots, d_i, \ldots)$ with $2d_i = \dim \mathrm{Ker}(((\mathrm{T}-\mathrm{a})^2 + \mathrm{b}^2)^{\mathrm{i}})$.

Corollary 1. If A, B are real matrices that are similar over $\mathbb{C}$ then they are similar over $\mathbb{R}$.

Corollary 2. Every real or complex square matrix is similar to its transpose.

Proof. Because of Corollary 1, we can assume that the matrix A is complex. We can also assume that A is in Jordan canonical form in which case we are reduced to proving the corollary in the case A is a Jordan matrix. This is left as an exercise for the reader.

If $f(X)$ is a polynomial over a field F, we can always construct a field K containing F as a subfield in which $f(X)$ splits completely into linear factors. One can even construct such a field which is algebraically closed. as a result, we obtain the Cayley-Hamilton Theorem. Namely, if $\Delta_T(X)$ is the characteristic polynomial of T, then $\Delta_T(T) = 0$.

There is a canonical form for the matrix of a linear operator T on a finite-dimensional vector space V over F, called the rational canonical form. The vector space V is a direct sum of cyclic subspaces, ie, subspaces having a basis of the form

$$v, T(v), \ldots, T^{p-1}(v)$$

with $T^p(v) = c_1 T^{p-1}(v) + \cdots + c_{p-1}T(v) + c_p v$. Such a vector v is called a cyclic vector for T. This subspace, called the cyclic subspace for T generated by v, is T-invariant and the matrix of the restriction of T to this subspace is the $p \times p$ matrix

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & c_p \\ 1 & 0 & 0 & \cdots & c_{p-1} \\ 0 & 1 & 0 & \cdots & c_{p-2} \\ 0 & 0 & 1 & \cdots & c_{p-3} \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & c_1 \end{bmatrix}.$$

This matrix is called the companion matrix of the cyclic vector v. Since the polynomial $f(X)$ is uniquely determined by v, this matrix is also called the companion matrix of $f(X)$ and is denoted by C_f. The polynomial $f(X)$ is the minimal polynomial of C_f and therefore of the restriction of T to the cyclic subspace generated by v.

If the minimal polynomial $m_T(X)$ of T has the primary decomposition

$$m_T(X) = p_1(X)_1^m \cdots p_k(X)_k^m$$

with $p_i(X)$ irreducible then V is the direct sum of cyclic subspaces with companion matrices C_f where $f(X) = p_i(X)^j$, $1 \leq i \leq k$, $1 \leq j \leq m_i$. Moreover, if s_j is the number of summands with $f(X) = p_i(X)^j$, then

$$s = (s_0, s_1, \ldots, s_j, \ldots) = -R(L-1)^2(d),$$

where $d = (d_0, d_1, \ldots, d_j, \ldots)$ with $d_j = \dim \mathrm{Ker}(\mathrm{p_i(T)^j})/\mathrm{degree}(\mathrm{p_i(X)}$. The proof of this result is exactly the same as the proof of the Jordan canonical form except that $p_i(T)$ replaces $T - a_i$ and the vector spaces $W_{ij} = \mathrm{Ker}(\mathrm{p_i(T)^j})/\mathrm{Ker}(\mathrm{p_i(T)^{j-1}}$ are viewed as vector spaces over the field $K_i = F[X]/(p_i(X))$, where

$$\overline{f(X)} \cdot \overline{g(T)} = \overline{f(T)g(T)}.$$

This field is a vector space over F of dimension $\ell_i = \mathrm{degree}(p_i(X))$ with basis $1, \overline{X}, \ldots, \overline{X}^{\ell_i - 1}$. Thus

$$\dim_F(W_{ij}) = \ell_i \dim_{K_i}(W_{ij}).$$

This explains the formula for d_j. The cyclic vectors giving the direct sum decomposition are representatives of basis vectors for the quotient spaces $W_{ij}/\mathrm{Im}(\mathrm{S_{ij}})$ where S_{ij} is the linear mapping from $W_{i(j+1)}$ to W_{ij} defined by

$$S_{ij}(\overline{u}) = \overline{p_i(T)(u)}.$$

The details are left to the reader.

It follows that two $n \times n$ matrices over F which are similar over a field containing F as a subfield are similar over F. This implies, for example, that a square matrix is similar to it transpose.

Example. Let us find the rational canonical form of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

over the field $\mathbb{Q}$ of rational numbers. This matrix can be viewed as an upper triangular 2×2 block matrix with blocks that are 2×2 matrices. Since the diagonal blocks have minimum polynomial $X^2 - 2$, it follows that $(A^2 - 2)^2 = 0$. We leave it to the reader to formulate a more general result. We have

$$A^2 - 2 = \begin{bmatrix} 2 & 0 & 2 & 3 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (A^2 - 2)^2 = 0$$

so that $\mathrm{Null(A^2 - 2) = Span(e_1, e_2)}$, $\mathrm{Null((A-2)^2) = (e_1, e_2, e_3, e_4)}$. Hence

$$\mathrm{Null((A^2-2)^2)/Null(A^2-2) = Span(\overline{e}_3, \overline{e}_4) = Span_K(\overline{e}_3)}$$

where $K = \mathbb{Q}[X]/(X^2 - 2) \cong \mathbb{Q}[\sqrt{2}]$. Note that $\overline{X}\overline{e}_3 = \overline{Ae_3} = 2\overline{e}_4$. We also have

$$\mathrm{Null(A^2 - 2) = Span(e_1, e_2) = Span_K(e_1)}$$

since $\overline{X}e_1 = Ae_1 = 2e_2$. Since $d = (0, 1, 2, 2, \ldots)$, we have $s = -R(L-1)^2(d) = (0, 0, 1, 0, 0, \ldots)$ so that we have only one cyclic subspace in the direct sum decomposition of $\mathbb{Q}^6$ with cyclic vector

e_3. Since $(X^2-2)^2 = X^4 - 4x^2 + 4$ is the companion polynomial of this cyclic vector, there is an invertible matrix over $\mathbb{Q}$ such that

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Moreover, the columns of P are the vectors $e_3, Ae_3, A^2e_3, A^3e_3$ so that

$$P = \begin{bmatrix} 0 & 1 & 2 & 4 \\ 0 & 1 & 0 & 4 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}.$$

The minimum of A is $(X^2-2)^2$. This is also the characteristic polynomial of A. We leave it to the reader to show that the Jordan canonical form of A over the field $\mathbb{Q}[\sqrt{2}]$ is

$$\begin{bmatrix} \sqrt{2} & 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & -\sqrt{2} & 1 \\ 0 & 0 & 0 & -\sqrt{2} \end{bmatrix}.$$