$$PAP^{-1} = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix}$$

where A_i is the companion matrix of the polynomial p_i. According to Theorem 7, the matrix

$$P(xI - A)P^{-1} = xI - PAP^{-1} \tag{7-33}$$

is equivalent to $xI - A$. Now

$$xI - PAP^{-1} = \begin{bmatrix} xI - A_1 & 0 & \cdots & 0 \\ 0 & xI - A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & xI - A_r \end{bmatrix} \tag{7-34}$$

where the various I's we have used are identity matrices of appropriate sizes. At the beginning of this section, we showed that $xI - A_i$ is equivalent to the matrix

$$\begin{bmatrix} p_i & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

From (7-33) and (7-34) it is then clear that $xI - A$ is equivalent to a diagonal matrix which has the polynomials p_i and $(n - r)$ 1's on its main diagonal. By a succession of row and column interchanges, we can arrange those diagonal entries in any order we choose, for example: $p_1, \ldots, p_r, 1, \ldots, 1$. ∎

Theorem 8 does not give us an effective way of calculating the elementary divisors $p_1, \ldots, p_r$ because our proof depends upon the cyclic decomposition theorem. We shall now give an explicit algorithm for reducing a polynomial matrix to diagonal form. Theorem 8 suggests that we may also arrange that successive elements on the main diagonal divide one another.

Definition. *Let* N *be a matrix in* $F[x]^{m \times n}$. *We say that* N *is in* (Smith) **normal form** *if*

(a) *every entry off the main diagonal of* N *is* 0;

(b) *on the main diagonal of* N *there appear (in order) polynomials* $f_1, \ldots, f_l$ *such that* f_k *divides* f_{k+1}, $1 \leq k \leq l - 1$.

In the definition, the number l is $l = \min(m, n)$. The main diagonal entries are $f_k = N_{kk}$, $k = 1, \ldots, l$.

Theorem 9. *Let* M *be an* $m \times n$ *matrix with entries in the polynomial algebra* $F[x]$. *Then* M *is equivalent to a matrix* N *which is in normal form.*

Proof. If $M = 0$, there is nothing to prove. If $M \neq 0$, we shall give an algorithm for finding a matrix M' which is equivalent to M and which has the form

$$M' = \begin{bmatrix} f_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & R & \\ 0 & & & \end{bmatrix} \tag{7-35}$$

where R is an $(m-1) \times (n-1)$ matrix and f_1 divides every entry of R. We shall then be finished, because we can apply the same procedure to R and obtain f_2, etc.

Let $l(M)$ be the minimum of the degrees of the non-zero entries of M. Find the first column which contains an entry with degree $l(M)$ and interchange that column with column 1. Call the resulting matrix $M^{(0)}$. We describe a procedure for finding a matrix of the form

$$\begin{bmatrix} g & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S & \\ 0 & & & \end{bmatrix} \tag{7-36}$$

which is equivalent to $M^{(0)}$. We begin by applying to the matrix $M^{(0)}$ the procedure of the lemma before Theorem 6, a procedure which we shall call PL6. There results a matrix

$$M^{(1)} = \begin{bmatrix} p & a & \cdots & b \\ 0 & c & \cdots & d \\ \vdots & \vdots & & \vdots \\ 0 & e & \cdots & f \end{bmatrix}. \tag{7-37}$$

If the entries $a, \ldots, b$ are all 0, fine. If not, we use the analogue of PL6 for the first row, a procedure which we might call PL6′. The result is a matrix

$$M^{(2)} = \begin{bmatrix} q & 0 & \cdots & 0 \\ a' & c' & \cdots & e' \\ \vdots & \vdots & & \vdots \\ b' & d' & \cdots & f' \end{bmatrix} \tag{7-38}$$

where q is the greatest common divisor of $p, a, \ldots, b$. In producing $M^{(2)}$, we may or may not have disturbed the nice form of column 1. If we did, we can apply PL6 once again. Here is the point. In not more than $l(M)$ steps:

$$M^{(0)} \xrightarrow{\text{PL6}} M^{(1)} \xrightarrow{\text{PL6}'} M^{(2)} \xrightarrow{\text{PL6}} \cdots \longrightarrow M^{(t)}$$

we must arrive at a matrix $M^{(t)}$ which has the form (7-36), because at each successive step we have $l(M^{(k+1)}) < l(M^{(k)})$. We name the process which we have just defined P7-36:

$$M^{(0)} \xrightarrow{\text{P7-36}} M^{(t)}.$$

In (7-36), the polynomial g may or may not divide every entry of S. If it does not, find the first column which has an entry not divisible by g and add that column to column 1. The new first column contains both g and an entry $gh + r$ where $r \neq 0$ and $\deg r < \deg g$. Apply process P7-36 and the result will be another matrix of the form (7-36), where the degree of the corresponding g has decreased.

It should now be obvious that in a finite number of steps we will obtain (7-35), i.e., we will reach a matrix of the form (7-36) where the degree of g cannot be further reduced. ∎

We want to show that the normal form associated with a matrix M is unique. Two things we have seen provide clues as to how the polynomials $f_1, \ldots, f_l$ in Theorem 9 are uniquely determined by M. First, elementary row and column operations do not change the determinant of a square matrix by more than a non-zero scalar factor. Second, elementary row and column operations do not change the greatest common divisor of the entries of a matrix.

Definition. *Let* M *be an* m $\times$ n *matrix with entries in* F[x]. *If* $1 \leq k \leq \min\,(m, n)$, *we define* $\delta_k(M)$ *to be the greatest common divisor of the determinants of all* k $\times$ k *submatrices of* M.

Recall that a $k \times k$ submatrix of M is one obtained by deleting some $m - k$ rows and some $n - k$ columns of M. In other words, we select certain k-tuples

$$\begin{aligned} I &= (i_1, \ldots, i_k), \qquad 1 \leq i_1 < \cdots < i_k \leq m \\ J &= (j_1, \ldots, j_k), \qquad 1 \leq j_1 < \cdots < j_k \leq n \end{aligned}$$

and look at the matrix formed using those rows and columns of M. We are interested in the determinants

$$D_{I,J}(M) = \det \begin{bmatrix} M_{i_1j_1} & \cdots & M_{i_1j_k} \\ \vdots & & \vdots \\ M_{i_kj_1} & \cdots & M_{i_kj_k} \end{bmatrix}. \tag{7-39}$$

The polynomial $\delta_k(M)$ is the greatest common divisor of the polynomials $D_{I,J}(M)$, as I and J range over the possible k-tuples.

Theorem 10. *If* M *and* N *are equivalent* m $\times$ n *matrices with entries in* F[x], *then*

$$\delta_k(M) = \delta_k(N), \qquad 1 \leq k \leq \min\,(m, n). \tag{7-40}$$

Proof. It will suffice to show that a single elementary row operation e does not change δ_k. Since the inverse of e is also an elementary row operation, it will suffice to show this: If a polynomial f divides every $D_{I,J}(M)$, then f divides $D_{I,J}(e(M))$ for all k-tuples I and J.

Since we are considering a row operation, let $\alpha_1, \ldots, \alpha_m$ be the rows of M and let us employ the notation

$$D_J(\alpha_{i_1}, \ldots, \alpha_{i_k}) = D_{I,J}(M).$$

Given I and J, what is the relation between $D_{I,J}(M)$ and $D_{I,J}(e(M))$? Consider the three types of operations e:

(a) multiplication of row r by a non-zero scalar c;
(b) replacement of row r by row r plus g times row s, $r \neq s$;
(c) interchange of rows r and s, $r \neq s$.

Forget about type (c) operations for the moment, and concentrate on types (a) and (b), which change only row r. If r is not one of the indices $i_1, \ldots, i_k$, then

$$D_{I,J}(e(M)) = D_{I,J}(M).$$

If r is among the indices $i_1, \ldots, i_k$, then in the two cases we have

$$\begin{aligned}
\text{(a)}\quad D_{I,J}(e(M)) &= D_J(\alpha_{i_1}, \ldots, c\alpha_r, \ldots, \alpha_{i_k})\\
&= cD_J(\alpha_{i_1}, \ldots, \alpha_r, \ldots, \alpha_{i_k})\\
&= cD_{I,J}(M);
\end{aligned}$$

$$\begin{aligned}
\text{(b)}\quad D_{I,J}(e(M)) &= D_J(\alpha_{i_1}, \ldots, \alpha_r + g\alpha_s, \ldots, \alpha_{i_k})\\
&= D_{I,J}(M) + gD_J(\alpha_{i_1}, \ldots, \alpha_s, \ldots, \alpha_{i_k}).
\end{aligned}$$

For type (a) operations, it is clear that any f which divides $D_{I,J}(M)$ also divides $D_{I,J}(e(M))$. For the case of a type (c) operation, notice that

$$\begin{aligned}
D_J(\alpha_{i_1}, \ldots, \alpha_s, \ldots, \alpha_{i_k}) &= 0, &&\text{if } s = i_j \text{ for some } j\\
D_J(\alpha_{i_1}, \ldots, \alpha_s, \ldots, \alpha_{i_k}) &= \pm D_{I',J}(M), &&\text{if } s \neq i_j \text{ for all } j.
\end{aligned}$$

The I' in the last equation is the k-tuple $(i_1, \ldots, s, \ldots, i_k)$ arranged in increasing order. It should now be apparent that, if f divides every $D_{I,J}(M)$, then f divides every $D_{I,J}(e(M))$.

Operations of type (c) can be taken care of by roughly the same argument or by using the fact that such an operation can be effected by a sequence of operations of types (a) and (b). ∎

Corollary. *Each matrix* M *in* $F[x]^{m\times n}$ *is equivalent to precisely one matrix* N *which is in normal form. The polynomials* $f_1, \ldots, f_l$ *which occur on the main diagonal of* N *are*

$$f_k = \frac{\delta_k(M)}{\delta_{k-1}(M)}, \qquad 1 \leq k \leq \min\,(m, n)$$

where, for convenience, we define $\delta_0(M) = 1$.

Proof. If N is in normal form with diagonal entries $f_1, \ldots, f_l$, it is quite easy to see that

$$\delta_k(N) = f_1 f_2 \cdots f_k. \quad ∎$$

Of course, we call the matrix N in the last corollary the **normal form** of M. The polynomials $f_1, \ldots, f_l$ are often called the **invariant factors** of M.

Suppose that A is an $n \times n$ matrix with entries in F, and let $p_1, \ldots, p_r$ be the invariant factors for A. We now see that the normal form of the matrix $xI - A$ has diagonal entries $1, 1, \ldots, 1, p_r, \ldots, p_1$. The last corollary tells us what $p_1, \ldots, p_r$ are, in terms of submatrices of $xI - A$. The number $n - r$ is the largest k such that $\delta_k(xI - A) = 1$. The minimal polynomial p_1 is the characteristic polynomial for A divided by the greatest common divisor of the determinants of all $(n - 1) \times (n - 1)$ submatrices of $xI - A$, etc.

Exercises

1. True or false? Every matrix in $F[x]^{n \times n}$ is row-equivalent to an upper-triangular matrix.

2. Let T be a linear operator on a finite-dimensional vector space and let A be the matrix of T in some ordered basis. Then T has a cyclic vector if and only if the determinants of the $(n - 1) \times (n - 1)$ submatrices of $xI - A$ are relatively prime.

3. Let A be an $n \times n$ matrix with entries in the field F and let $f_1, \ldots, f_n$ be the diagonal entries of the normal form of $xI - A$. For which matrices A is $f_1 \neq 1$?

4. Construct a linear operator T with minimal polynomial $x^2(x - 1)^2$ and characteristic polynomial $x^3(x - 1)^4$. Describe the primary decomposition of the vector space under T and find the projections on the primary components. Find a basis in which the matrix of T is in Jordan form. Also find an explicit direct sum decomposition of the space into T-cyclic subspaces as in Theorem 3 and give the invariant factors.

5. Let T be the linear operator on R^8 which is represented in the standard basis by the matrix

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

(a) Find the characteristic polynomial and the invariant factors.

(b) Find the primary decomposition of R^8 under T and the projections on the primary components. Find cyclic decompositions of each primary component as in Theorem 3.