Define $d_1, \dots, d_m$ to be the degrees of the monic nonconstant polynomials $a_1(x), \dots, a_m(x)$ appearing on the diagonal, respectively.

(2) Beginning with the $n \times n$ identity matrix P', for each row operation used in (1), change the matrix P' by the following rules:

(a) If $R_i \leftrightarrow R_j$ then interchange the i^{th} and j^{th} columns of P' (i.e., $C_i \leftrightarrow C_j$ for P').

(b) If $R_i + p(x)R_j \mapsto R_i$ then subtract the product of the matrix $p(A)$ times the i^{th} column of P' from the j^{th} column of P' (i.e., $C_j - p(A)C_i \mapsto C_j$ for P' — note the indices).

(c) If uR_i then divide the elements of the i^{th} column of P' by u (i.e., $u^{-1}C_i$ for P').

(3) When $xI - A$ has been diagonalized to the form in Theorem 21 the first $n - m$ columns of the matrix P' are 0 (providing a useful numerical check on the computations) and the remaining m columns of P' are nonzero. For each $i = 1, 2, \dots, m$, multiply the i^{th} nonzero column of P' successively by $A^0 = I, A^1, A^2, \dots, A^{d_i - 1}$, where d_i is the integer in (1) above and use the resulting column vectors (in this order) as the next d_i columns of an $n \times n$ matrix P. Then $P^{-1}AP$ is in rational canonical form (whose diagonal blocks are the companion matrices for the polynomials $a_1(x), \dots, a_m(x)$ in (1)).

In the theory of canonical forms for linear transformations (or matrices) the characteristic polynomial plays the role of the order of a finite abelian group and the minimal polynomial plays the role of the exponent (after all, they are the same invariants, one for modules over the Principal Ideal Domain $\mathbb{Z}$ and the other for modules over the Principal Ideal Domain $F[x]$) so we can solve problems directly analogous to those we considered for finite abelian groups in Chapter 5. In particular, this includes the following:

(A) determine the rational canonical form of a given matrix (analogous to decomposing a finite abelian group as a direct product of cyclic groups)

(B) determine whether two given matrices are similar (analogous to determining whether two given finite abelian groups are isomorphic)

(C) determine all similarity classes of matrices over F with a given characteristic polynomial (analogous to determining all abelian groups of a given order)

(D) determine all similarity classes of $n \times n$ matrices over F with a given minimal polynomial (analogous to determining all abelian groups of rank at most n of a given exponent).

Examples

(1) We find the rational canonical forms of the following matrices over $\mathbb{Q}$ and determine if they are similar:

$$A = \begin{pmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 0 & -4 & 85 \\ 1 & 4 & -30 \\ 0 & 0 & 3 \end{pmatrix} \quad C = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{pmatrix}.$$

A direct computation shows that all three of these matrices have the same characteristic polynomial: $c_A(x) = c_B(x) = c_C(x) = (x-2)^2(x-3)$. Since the minimal and characteristic polynomials have the same roots, the only possibilities for the minimal polynomials are $(x-2)(x-3)$ or $(x-2)^2(x-3)$. We quickly find that $(A-2I)(A-3I)=0$, $(B-2I)(B-3I)\neq 0$ (the 1,1-entry is nonzero) and $(C-2I)(C-3I)\neq 0$ (the 1,2-entry is nonzero). It follows that

$$m_A(x)=(x-2)(x-3), \quad m_B(x)=m_C(x)=(x-2)^2(x-3).$$

It follows immediately that there are no additional invariant factors for B and C. Since the invariant factors for A divide the minimal polynomial and have product the characteristic polynomial, we see that A has for invariant factors the polynomials $x-2$, $(x-2)(x-3)=x^2-5x+6$. (For 2×2 and 3×3 matrices the determination of the characteristic and minimal polynomials determines all the invariant factors, cf. Exercises 3 and 4.) We conclude that B and C are similar and neither is similar to A. The rational canonical forms are (note $(x-2)^2(x-3)=x^3-7x^2+16x-12$)

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 12 \\ 1 & 0 & -16 \\ 0 & 1 & 7 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 12 \\ 1 & 0 & -16 \\ 0 & 1 & 7 \end{pmatrix}.$$

(2) In the example above the rational canonical forms were obtained simply by determining the characteristic and minimal polynomials for the matrices. As mentioned, this is sufficient for 2×2 and 3×3 matrices since this information is sufficient to determine all of the invariant factors. For larger matrices, however, this is in general not sufficient (cf. the next example) and more work is required to determine the invariant factors. In this example we again compute the rational canonical form for the matrix A in Example 1 following the two algorithms outlined above. While this is computationally more difficult for this small matrix (as will be apparent), it has the advantage even in this case that it also explicitly computes a matrix P with $P^{-1}AP$ in rational canonical form.

I. (*Invariant Factor Decomposition*) We use row and column operations (in $\mathbb{Q}[x]$) to reduce the matrix

$$xI-A=\begin{pmatrix} x-2 & 2 & -14 \\ 0 & x-3 & 7 \\ 0 & 0 & x-2 \end{pmatrix}$$

to diagonal form. As in the invariant factor decomposition algorithm, we shall use the notation $R_i \leftrightarrow R_j$ to denote the interchange of the i^{th} and j^{th} rows, $R_i+aR_j \mapsto R_i$ if a times the j^{th} row is added to the i^{th} row, simply uR_i if the i^{th} row is multiplied by u (and similarly for columns, using C instead of R). Note also that the first two operations we perform below are rather *ad hoc* and were chosen simply to have integers everywhere in the computation:

$$\begin{pmatrix} x-2 & 2 & -14 \\ 0 & x-3 & 7 \\ 0 & 0 & x-2 \end{pmatrix} \xrightarrow[\mapsto R_1]{R_1+R_2} \begin{pmatrix} x-2 & x-1 & -7 \\ 0 & x-3 & 7 \\ 0 & 0 & x-2 \end{pmatrix} \longrightarrow$$

$$\xrightarrow[\mapsto C_1]{C_1-C_2} \begin{pmatrix} -1 & x-1 & -7 \\ -x+3 & x-3 & 7 \\ 0 & 0 & x-2 \end{pmatrix} \xrightarrow[-R_1]{} \begin{pmatrix} 1 & -x+1 & 7 \\ -x+3 & x-3 & 7 \\ 0 & 0 & x-2 \end{pmatrix} \longrightarrow$$

$$\xrightarrow[\substack{R_2+(x-3)R_1\\ \mapsto R_2}]{} \begin{pmatrix} 1 & -x+1 & 7 \\ 0 & -x^2+5x-6 & 7(x-2) \\ 0 & 0 & x-2 \end{pmatrix} \xrightarrow[\substack{C_2+(x-1)C_1\\ \mapsto C_2}]{} \begin{pmatrix} 1 & 0 & 7 \\ 0 & -x^2+5x-6 & 7(x-2) \\ 0 & 0 & x-2 \end{pmatrix} \longrightarrow$$

$$\xrightarrow[\substack{C_3-7C_1\\ \mapsto C_3}]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -x^2+5x-6 & 7(x-2) \\ 0 & 0 & x-2 \end{pmatrix} \xrightarrow[-C_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x^2-5x+6 & 7(x-2) \\ 0 & 0 & x-2 \end{pmatrix} \longrightarrow$$

$$\xrightarrow[\substack{R_2-7R_3\\ \mapsto R_2}]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x^2-5x+6 & 0 \\ 0 & 0 & x-2 \end{pmatrix} \xrightarrow[\substack{R_2\leftrightarrow R_3\\ C_2\leftrightarrow C_3}]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-2 & 0 \\ 0 & 0 & x^2-5x+6 \end{pmatrix}.$$

This determines the invariant factors $x-2, x^2-5x+6$ for this matrix, which we determined in Example 1 above. Let now V be a 3-dimensional vector space over $\mathbb{Q}$ with basis e_1, e_2, e_3 and let T be the corresponding linear transformation (which defines the action of x on V), i.e.,

$$\begin{aligned} xe_1 &= T(e_1) = 2e_1 \\ xe_2 &= T(e_2) = -2e_1 + 3e_2 \\ xe_3 &= T(e_3) = 14e_1 - 7e_2 + 2e_3. \end{aligned}$$

The row operations used in the reduction above were

$$R_1 + R_2 \mapsto R_1,\ -R_1,\ R_2 + (x-3)R_1 \mapsto R_2,\ R_2 - 7R_3 \mapsto R_2,\ R_2 \leftrightarrow R_3.$$

Starting with the basis $[e_1, e_2, e_3]$ for V and changing it according to the rules given in the text, we obtain

$$\begin{aligned} [e_1, e_2, e_3] &\longrightarrow [e_1, e_2-e_1, e_3] \longrightarrow [-e_1, e_2-e_1, e_3] \\ &\longrightarrow [-e_1-(x-3)(e_2-e_1), e_2-e_1, e_3] \\ &\longrightarrow [-e_1-(x-3)(e_2-e_1), e_2-e_1, e_3+7(e_2-e_1)] \\ &\longrightarrow [-e_1-(x-3)(e_2-e_1), e_3+7(e_2-e_1), e_2-e_1]. \end{aligned}$$

Using the formulas above for the action of x, we see that these last elements are the elements $[0, -7e_1 + 7e_2 + e_3, -e_1 + e_2]$ of V corresponding to the elements $1, x-2$ and x^2-5x+6 in the diagonalized form of $xI - A$, respectively. The elements $f_1 = -7e_1 + 7e_2 + e_3$ and $f_2 = -e_1 + e_2$ are therefore $\mathbb{Q}[x]$-module generators for the two cyclic factors of V in its invariant factor decomposition as a $\mathbb{Q}[x]$-module. The corresponding $\mathbb{Q}$-vector space bases for these two factors are then f_1 and $f_2, xf_2 = Tf_2$, i.e., $-7e_1+7e_2+e_3$ and $-e_1+e_2, T(-e_1+e_2) = -4e_1+3e_2$. Then the matrix

$$P = \begin{pmatrix} -7 & -1 & -4 \\ 7 & 1 & 3 \\ 1 & 0 & 0 \end{pmatrix}$$

conjugates A into its rational canonical form:

$$P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix},$$

as one easily checks.

II. (*Converting A Directly to Rational Canonical Form*) We use the row operations involved in the diagonalization of $xI - A$ to determine the matrix P' of the algorithm above:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underset{\substack{C_2-C_1 \\ \mapsto C_2}}{\longrightarrow} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underset{-C_1}{\longrightarrow} \begin{pmatrix} -1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow$$

$$\underset{\substack{C_1-(A-3I)C_2 \\ \mapsto C_1}}{\longrightarrow} \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underset{\substack{C_3+7C_2 \\ \mapsto C_3}}{\longrightarrow} \begin{pmatrix} 0 & -1 & -7 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix} \underset{C_2 \leftrightarrow C_3}{\longrightarrow} \begin{pmatrix} 0 & -7 & -1 \\ 0 & 7 & 1 \\ 0 & 1 & 0 \end{pmatrix} = P'.$$

Here we have $d_1 = 1$ and $d_2 = 2$, corresponding to the second and third nonzero columns of P', respectively. The columns of P are therefore given by

$$\begin{pmatrix} -7 \\ 7 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad A\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix},$$

respectively, which again gives the matrix P above.

(3) For the 3×3 matrix A it was not necessary to perform the lengthy calculations above merely to determine the rational canonical form (equivalently, the invariant factors), as we saw in Example 1. For $n \times n$ matrices with $n \geq 4$, however, the computation of the characteristic and minimal polynomials is in general not sufficient for the determination of all the invariant factors, so the more extensive calculations of the previous example may become necessary. For example, consider the matrix

$$D = \begin{pmatrix} 1 & 2 & -4 & 4 \\ 2 & -1 & 4 & -8 \\ 1 & 0 & 1 & -2 \\ 0 & 1 & -2 & 3 \end{pmatrix}.$$

A short computation shows that the characteristic polynomial of D is $(x-1)^4$. The possible minimal polynomials are then $x-1$, $(x-1)^2$, $(x-1)^3$ and $(x-1)^4$. Clearly $D - I \neq 0$ and another short computation shows that $(D-I)^2 = 0$, so the minimal polynomial for D is $(x-1)^2$. There are then two possible sets of invariant factors:

$$x-1, x-1, (x-1)^2 \qquad \text{and} \qquad (x-1)^2, (x-1)^2.$$

To determine the invariant factors for D we apply the procedure of the previous example to the 4×4 matrix

$$xI - D = \begin{pmatrix} x-1 & -2 & 4 & -4 \\ -2 & x+1 & -4 & 8 \\ -1 & 0 & x-1 & 2 \\ 0 & -1 & 2 & x-3 \end{pmatrix}.$$

The diagonal matrix obtained from this matrix by elementary row and column operations is the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (x-1)^2 & 0 \\ 0 & 0 & 0 & (x-1)^2 \end{pmatrix},$$

which shows that the invariant factors for D are $(x-1)^2$, $(x-1)^2$ (one series of elementary row and column operations which diagonalize $xI - D$ are $R_1 \leftrightarrow R_3$, $-R_1$,