A CRASH COURSE
ON THE CANONICAL FORM OF A MATRIX

Alfonso Gracia–Saz

Disclaimer: These notes have not been properly edited. If you find any mistakes, please let me know. For proofs and details, see any linear algebra textbook. Dummit and Foote also covers this material, in Chapter 12, but in much more generality.

Let $\mathbb{F}$ be a field. Let n be an integer. We denote by $\mathrm{Mat}(n,\mathbb{F})$ the set of square n by n matrices with coefficients in $\mathbb{F}$. We denote by $\mathrm{GL}(n,\mathbb{F})$ the set of such matrices with non-zero determinant. All polynomials in this paper have coefficients in $\mathbb{F}$.

The group $\mathrm{GL}(n,\mathbb{F})$ acts on the set $\mathrm{Mat}(n,\mathbb{F})$ by conjugation. Let $M,N \in \mathrm{Mat}(n,\mathbb{F})$. We say that M and N are *conjugate* when $N = AMA^{-1}$ for some $A \in \mathrm{GL}(n,\mathbb{F})$. Our goal is to classify matrices up to conjugation.

1 The characteristic and the minimal polynomial of a matrix

Let A be an $n \times n$ matrix. We associate two polynomials to A:

1. The *characteristic polynomial* of A is defined as $f(X) = \det(X \cdot 1 - A)$, where X is the variable of the polynomial, and 1 represents the identity matrix. $f(X)$ is a monic polynomial of degree n.

2. The *minimal polynomial* of A, which we will denote by $\mu(X)$, is defined by the following properties:

   - $\mu(X)$ is monic (i.e., its leading coefficent is 1),
   - $\mu(A) = 0$,
   - $\mu(X)$ is the monic polynomial of the smallest possible degree such that $\mu(A) = 0$,

They also satisfy the following properties:

- If $g(X)$ is another polynomial, then $g(A) = 0$ if and only if $\mu(X)$ divides $g(X)$.
- $f(X)$ is a multiple of $\mu(X)$.

For instance, the characteristic and the minimal polynomials of the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are, respectively, $X^3 - X^2 - X + 1$ and $X^2 - 1$.

2 The elementary divisors of a matrix

To every matrix $A \in \mathrm{Mat}(n, \mathbb{F})$ we can associate a set of polynomials $p_1(X), \ldots, p_r(X)$, called the *elementary divisors of* A. The number of elementary divisors depends on the matrix. I am not going to give you the definition, but here are some properties.

- Each elementary divisor $p_i(X)$ is a monic polynomial.
- Each elementary divisor $p_i(X)$ is a power of an irreducible polynomial (i.e., one that cannot be factored).
- The characteristic polynomial of A is the product of all the elementary divisors. Hence, the sum of the degrees of the minimal polynomials equals the size of A.
- The minimal polynomial of A is the least common multiple of all the elementary divisors.

Knowing the minimal polynomial of a matrix often leaves only a few possibilities for the elementary divisors. For instance, let $A \in \mathrm{Mat}(6, \mathbb{R})$. Assume that the minimal polynomial of A is $(X^2 + 1)(X + 1)^2$. You should be able to compute that the only options for the elementary divisors of A are:

- $(X^2 + 1), (X^2 + 1), (X + 1)^2$;
- $(X^2 + 1), (X + 1)^2, (X + 1)^2$;
- $(X^2 + 1), (X + 1)^2, X + 1, X + 1$.

3 The accompanying matrix of a set of polynomials

Given a monic polynomial $p(X) = X^m + a_{m-1}X^{m-1} + \ldots + a_1 x + a_0$, we define its accompanying matrix as

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{m-1} \end{pmatrix}$$

Notice that this is a square matrix with size equal to the degree of $p(X)$.

Given a collection of monic polynomials $p_1(X), \ldots, p_r(X)$, its accompanying matrix is the block–diagonal matrix, whose diagonal blocks are the accompanying matrices of $p_1(X), \ldots, p_r(X)$.

For instance, the accompanying matrix of $\{X^2 + 1, X^3 - 2X^2 + 7\}$ is

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

where

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 & -7 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

4 The main result

Two matrices are conjugate if and only if they have the same elementary divisors.

In addition, each conjugacy class has one special representative. Given a set of elementary divisors, the accompanying matrix of that set has exactly that set as elementary divisors.

5 An application

We want to find all matrices in $\mathrm{GL}(3, \mathbb{F}_2)$ which have order 7, where $\mathbb{F}_2$ is the field with 2 elements.

Let A be one such matrix. A satisfies $A^7 - 1 = 0$. Let $\mu(A)$ be the minimal polynomial of A. We know that $\mu(A)$, has to be a divisor of $X^7 - 1$. Notice that in this field $1 = -1$. We factor this polynomial as product of irreducibles:

$$X^7 + 1 = (X + 1)(X^3 + X^2 + 1)(X^3 + X + 1)$$

Since $\mu(A)$ has degree at most 3, there are only three options:

1. $\mu(A) = X^3 + X^2 + 1$. In this case A has only one elementary divisor, namely $\mu(A)$.
2. $\mu(A) = X^3 + X + 1$. In this case A has only one elementary divisor, namely $\mu(A)$.
3. $\mu(A) = X + 1$, but this corresponds to the identity matrix, which does not have order 7.

One example of matrix in cases 1 and 2 will be the accompanying matrix of each of those polynomials:

$$M_1 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

You can check directly that M_1 and M_2 indeed have order 7.

We conclude that the matrices in $\mathrm{GL}(3, \mathbb{F}_2)$ which have order 7 are exactly the matrices which are conjugate to either M_1 or M_2.