

## 6 The Smith Canonical Form

### 6.1 Equivalence of Polynomial Matrices

#### DEFINITION 6.1

A matrix  $P \in M_{n \times n}(F[x])$  is called a **unit** in  $M_{n \times n}(F[x])$  if  $\exists Q \in M_{n \times n}(F[x])$  such that

$$PQ = I_n.$$

Clearly if  $P$  and  $Q$  are units, so is  $PQ$ .

#### THEOREM 6.1

A matrix  $P \in M_{n \times n}(F[x])$  is a unit in  $M_{n \times n}(F[x])$  if and only if  $\det P = c$ , where  $c \in F$  and  $c \neq 0$ .

proof

“only if”. Suppose  $P$  is a unit. Then  $PQ = I_n$  and

$$\det PQ = \det P \det Q = \det I_n = 1.$$

However  $\det P$  and  $\det Q$  belong to  $F[x]$ , so both are in fact non-zero elements of  $F$ .

“if”. Suppose  $P \in M_{n \times n}(F[x])$  satisfies  $\det P = c$ , where  $c \in F$  and  $c \neq 0$ . Then

$$P \operatorname{adj} P = (\det P)I_n = cI_n.$$

Hence  $PQ = I_n$ , where  $Q = c^{-1} \operatorname{adj} P \in M_{n \times n}(F[x])$ . Hence  $P$  is a unit in  $M_{n \times n}(F[x])$ .

#### EXAMPLE 6.1

$$P = \begin{bmatrix} 1+x & -x \\ x & 1-x \end{bmatrix} \in M_{2 \times 2}(F[x]) \text{ is a unit, as } \det P = 1.$$

#### THEOREM 6.2

Elementary row matrices in  $M_{n \times n}(F[x])$  are units:

- (i)  $E_{ij}$ : interchange rows  $i$  and  $j$  of  $I_n$ ;
- (ii)  $E_i(t)$ : multiply row  $i$  of  $I_n$  by  $t \in F$ ,  $t \neq 0$ ;
- (iii)  $E_{ij}(f)$ : add  $f$  times row  $j$  of  $I_n$  to row  $i$ ,  $f \in F[x]$ .

In fact  $\det E_{ij} = -1$ ;  $\det E_i(t) = t$ ;  $\det E_{ij}(f) = 1$ .

Similarly for elementary column matrices in  $M_{n \times n}(F[x])$ :

$$F_{ij}, F_i(t), F_{ij}(f).$$

REMARK: It follows that a product of elementary matrices in  $M_{n \times n}(F[x])$  is a unit. Later we will be able to prove that the converse is also true.

**DEFINITION 6.2**

Let  $A, B \in M_{m \times n}(F[x])$ . Then  $A$  is equivalent to  $B$  over  $F[x]$  if units  $P \in M_{m \times m}(F[x])$  and  $Q \in M_{n \times n}(F[x])$  exist such that

$$PAQ = B.$$

**THEOREM 6.3**

Equivalence of matrices over  $F[x]$  defines an equivalence relation on  $M_{m \times n}(F[x])$ .

**6.1.1 Determinantal Divisors**

**DEFINITIONS 6.1**

Let  $A \in M_{m \times n}(F[x])$ . Then for  $1 \leq k \leq \min(m, n)$ , let  $d_k(A)$  denote the gcd of all  $k \times k$  minors of  $A$ .

$d_k(A)$  is sometimes called the  $k^{\text{th}}$  **determinantal divisor** of  $A$ .

**Note:**  $\gcd(f_1, \dots, f_n) \neq 0 \Leftrightarrow$  at least one of  $f_1, \dots, f_n$  is non-zero.

$\rho(A)$ , the **determinantal rank** of  $A$ , is defined to be the largest integer  $r$  for which there exists a non-zero  $r \times r$  minor of  $A$ .

**THEOREM 6.4**

For  $1 \leq k \leq \rho(A)$ , we have  $d_k(A) \neq 0$ . Also  $d_k(A)$  divides  $d_{k+1}(A)$  for  $1 \leq k \leq \rho(A) - 1$ .

proof

Let  $r = \rho(A)$ . Then there exists an  $r \times r$  non-zero minor and hence  $d_r(A) \neq 0$ . Then because each  $r \times r$  minor is a linear combination over  $F[x]$  of  $(r-1) \times (r-1)$  minors of  $A$ , it follows that some  $(r-1) \times (r-1)$  minor of  $A$  is also non-zero and hence  $d_{r-1}(A) \neq 0$ ; also  $d_{r-1}(A)$  divides each minor of size  $r-1$  and consequently divides each minor of size  $r$ ; hence  $d_{r-1}(A)$  divides  $d_r(A)$ , the gcd of all minors of size  $r$ . This argument can be repeated with  $r$  replaced by  $r-1$  and so on.

**THEOREM 6.5**

Let  $A, B \in M_{m \times n}(F[x])$ . Then if  $A$  is equivalent to  $B$  over  $F[x]$ , we have

(i)  $\rho(A) = \rho(B) = r$ ;

(ii)  $d_k(A) = d_k(B)$  for  $1 \leq k \leq r$ .

proof

Suppose  $PAQ = B$ , where  $P$  and  $Q$  are units. First consider  $PA$ . The rows of  $PA$  are linear combinations over  $F[x]$  of the rows of  $A$ , so it follows that each  $k \times k$  minor of  $PA$  is a linear combination of the  $k \times k$  minors of  $A$ . Similarly each column of  $(PA)Q$  is a linear combinations over  $F[x]$  of the columns of  $PA$ , so it follows that each  $k \times k$  minor of  $B = (PA)Q$  is a linear combination over  $F[x]$  of the  $k \times k$  minors of  $PA$  and consequently of the  $k \times k$  minors of  $A$ .

It follows that all minors of  $B$  with size  $k > \rho(A)$  must be zero and hence  $\rho(B) \leq \rho(A)$ . However  $B$  is equivalent to  $A$ , so we deduce that  $\rho(A) \leq \rho(B)$  and hence  $\rho(A) = \rho(B)$ .

Also  $d_k(B)$  is a linear combination over  $F[x]$  of all  $k \times k$  minors of  $B$  and hence of all  $k \times k$  minors of  $A$ . Hence  $d_k(A) | d_k(B)$  and by symmetry,  $d_k(B) | d_k(A)$ . Hence  $d_k(A) = d_k(B)$  if  $1 \leq k \leq r$ .

## 6.2 Smith Canonical Form

### THEOREM 6.6 (Smith canonical form)

Every non-zero matrix  $A \in M_{m \times n}(F[x])$  with  $r = \rho(A)$  is equivalent to a matrix of the form

$$D = \begin{bmatrix} f_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & f_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & f_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} = PAQ$$

where  $f_1, \dots, f_r \in F[x]$  are monic,  $f_k | f_{k+1}$  for  $1 \leq k \leq r-1$ ,  $P$  is a product of elementary row matrices, and  $Q$  is a product of elementary column matrices.

### DEFINITION 6.3

The matrix  $D$  is said to be in **Smith canonical form**.

proof

This is presented in the form of an algorithm which is in fact used by CMAT to find unit matrices  $P$  and  $Q$  such that  $PAQ$  is in Smith canonical form.

Our account is based on that in the book “Rings, Modules and Linear Algebra,” by B. Hartley and T.O. Hawkes.

We describe a sequence of elementary row and column operations over  $F[x]$ , which when applied to a matrix  $A$  with  $a_{11} \neq 0$  either yields a matrix  $C$  of the form

$$C = \begin{bmatrix} f_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & C^* & \\ 0 & & & \end{bmatrix}$$

where  $f_1$  is monic and divides every element of  $C^*$ , or else yields a matrix  $B$  in which  $b_{11} \neq 0$  and

$$\deg b_{11} < \deg a_{11}. \quad (28)$$

Assuming this, we start with our non-zero matrix  $A$ . By performing suitable row and column interchanges, we can assume that  $a_{11} \neq 0$ . Now repeatedly perform the algorithm mentioned above. Eventually we must reach a matrix of type  $C$ , otherwise we would produce an infinite strictly decreasing sequence of non-negative integers by virtue of inequalities of type (28).

On reaching a matrix of type  $C$ , we stop if  $C^* = 0$ . Otherwise we perform the above argument on  $C^*$  and so on, leaving a trail of diagonal elements as we go.

Two points must be made:

- (i) Any elementary row or column operation on  $C^*$  corresponds to an elementary operation on  $C$ , which does not affect the first row or column of  $C$ .
- (ii) Any elementary operation on  $C^*$  gives a new  $C^*$  whose new entries are linear combinations over  $F[x]$  of the old ones; consequently these new entries will still be divisible by  $f_1$ .

Hence in due course we will reach a matrix  $D$  which is in Smith canonical form.

We now detail the sequence of elementary operations mentioned above. Case 1.  $\exists a_{1j}$  in row 1 with  $a_{11}$  not dividing  $a_{1j}$ . Then

$$a_{1j} = a_{11}q + b,$$

by Euclid’s division theorem, where  $b \neq 0$  and  $\deg b < \deg a_{11}$ . Subtract  $q$  times column 1 from column  $j$  and then interchange columns 1 and  $j$ . This yields a matrix of type  $B$  mentioned above.

Case 2.  $\exists a_{i1}$  in column 1 with  $a_{11}$  not dividing  $a_{i1}$ . Proceed as in Case 1, operating on rows rather than columns, again reaching a matrix of type  $B$ .  
 Case 3. Here  $a_{11}$  divides every element in the first row and first column. Then by subtracting suitable multiples of column 1 from the other columns, we can replace all the entries in the first row other than  $a_{11}$  by 0. Similarly for the first column. We then have a matrix of the form

$$E = \begin{bmatrix} e_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & E^* & \\ 0 & & & \end{bmatrix}.$$

If  $e_{11}$  divides every element of  $E^*$ , we have reached a matrix of type  $C$ . Otherwise  $\exists e_{ij}$  not divisible by  $e_{11}$ . We then add row  $i$  to row 1, thereby reaching Case 1.

**EXAMPLE 6.2**

(of the Smith Canonical Form)

$$A = \begin{bmatrix} 1+x^2 & x \\ x & 1+x \end{bmatrix}$$

We want  $D = PAQ$  in Smith canonical form. So we construct the augmented matrix

	work on rows	work on columns
	↓	↓
	1    0	1+x <sup>2</sup> x
	0    1	x    1+x
$R_1 \rightarrow R_1 - xR_2 \Rightarrow$	1    -x	1    -x <sup>2</sup>
	0    1	x    1+x
$C_2 \rightarrow C_2 + x^2C_1 \Rightarrow$	1    -x	1    0
	0    1	x    1+x+x <sup>3</sup>
$R_2 \rightarrow R_2 - xR_1 \Rightarrow$	1    -x	1    0
	-x   1+x <sup>2</sup>	0    1+x+x <sup>3</sup>
	↑	↑
	$P$	$D$
		↑
		$Q$

Invariants are  $f_1 = 1$ ,  $f_2 = 1 + x + x^3$ . Note also

$$f_1 = d_1(A), \quad f_2 = \frac{d_2(A)}{d_1(A)}.$$

### 6.2.1 Uniqueness of the Smith Canonical Form

#### THEOREM 6.7

Every matrix  $A \in M_{m \times n}(F[x])$  is equivalent to precisely one matrix in Smith canonical form.

proof Suppose  $A$  is equivalent to a matrix  $B$  in Smith canonical form. That is,

$$B = \left[ \begin{array}{ccc|c} f_1 & & & 0 \\ & \ddots & & \\ & & f_r & \\ \hline & & & 0 \end{array} \right] \quad \text{and} \quad f_1 \mid f_2 \mid \cdots \mid f_r.$$

Then  $r = \rho(A)$ , the determinantal rank of  $A$ . But if  $1 \leq k \leq r$ ,

$$d_k(A) = d_k(B) = f_1 f_2 \cdots f_k$$

and so the  $f_i$  are uniquely determined by

$$\begin{aligned} f_1 &= d_1(A) \\ f_2 &= \frac{d_2(A)}{d_1(A)} \\ &\vdots \\ f_r &= \frac{d_r(A)}{d_{r-1}(A)}. \end{aligned}$$

### 6.3 Invariant factors of a polynomial matrix

#### DEFINITION 6.4

The polynomials  $f_1, \dots, f_r$  in the Smith canonical form of  $A$  are called the **invariant factors** of  $A$ .<sup>3</sup>

**Note:** CMAT calls the invariant factors of  $xI - B$ , where  $B \in M_{n \times n}(F)$ , the “similarity invariants” of  $B$ .

We next find these similarity invariants. They are

$$\underbrace{1, 1, \dots, 1}_{n-s}, d_1, \dots, d_s$$

where  $d_1, \dots, d_s$  are what earlier called the invariant factors of  $T_B$ .

<sup>3</sup>**NB.** This is a slightly different, though similar, form of “invariant factor” to that we met a short while ago.

**LEMMA 6.1**

The Smith canonical form of  $xI_n - C(d)$  where  $d$  is a monic polynomial of degree  $n$  is

$$\text{diag}(\underbrace{1, \dots, 1}_{n-1}, d).$$

proof Let  $d = x^n + a_{n-1}x^{n-1} + \dots + a_0 \in F[x]$ , so

$$xI_n - C(d) = \begin{bmatrix} x & 0 & & & a_0 \\ -1 & x & \cdots & & a_1 \\ 0 & -1 & & & a_2 \\ & \vdots & \ddots & & \vdots \\ & & & x & a_{n-2} \\ 0 & & \cdots & -1 & x + a_{n-1} \end{bmatrix}.$$

Now use the row operation

$$R_1 \rightarrow R_1 + xR_2 + x^2R_3 + \dots + x^{n-1}R_n$$

to obtain

$$\begin{bmatrix} 0 & 0 & & & d \\ -1 & x & \cdots & & a_1 \\ 0 & -1 & & & a_2 \\ & \vdots & \ddots & & \vdots \\ & & & x & a_{n-2} \\ 0 & & \cdots & -1 & x + a_{n-1} \end{bmatrix}$$

(think about it!) and then column operations

$$C_2 \rightarrow C_2 + xC_1, \dots, C_{n-1} \rightarrow C_{n-1} + xC_{n-2}$$

and then

$$C_n \rightarrow C_n + a_1C_1 + a_2C_2 + \dots + a_{n-2}C_{n-2} + (x + a_{n-1})C_{n-1}$$

yielding

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & d \\ -1 & 0 & & & 0 \\ 0 & -1 & & & \\ & & \ddots & & \vdots \\ 0 & & \cdots & -1 & 0 \end{bmatrix}.$$



by finding the Smith canonical form of  $xI_4 - B$ .

**Solution:**

$$xI_4 - B = \begin{bmatrix} x-2 & 0 & 0 & 0 \\ 1 & x-1 & 0 & 0 \\ 0 & 1 & x & 1 \\ -1 & -1 & -1 & x-2 \end{bmatrix}$$

We start off with the row operations

$$\begin{aligned} R_1 &\rightarrow R_1 - (x-2)R_2 \\ R_1 &\leftrightarrow R_2 \\ R_4 &\rightarrow R_4 + R_1 \end{aligned}$$

and get

$$\begin{aligned} &\begin{bmatrix} 1 & x-1 & 0 & 0 \\ 0 & -(x-1)(x-2) & 0 & 0 \\ 0 & 1 & x & 1 \\ 0 & x-2 & -1 & x-2 \end{bmatrix} \\ \text{(column ops.) } \Rightarrow &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \overline{-(x-1)(x-2)} & 0 & 0 \\ 0 & 1 & x & 1 \\ 0 & x-2 & -1 & x-2 \end{bmatrix} \\ \Rightarrow &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & 1 \\ 0 & -(x-1)(x-2) & 0 & 0 \\ 0 & x-2 & -1 & x-2 \end{bmatrix} \\ \Rightarrow &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & x(x-1)(x-2) & (x-1)(x-2) \\ 0 & 0 & -1-x(x-2) & 0 \\ & & \{= -(x-1)^2\} & \end{bmatrix} \\ \Rightarrow &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x(x-1)(x-2) & (x-1)(x-2) \\ 0 & 0 & -(x-1)^2 & 0 \end{bmatrix}. \end{aligned}$$

Now, for brevity, we work just on the  $2 \times 2$  block in the bottom right corner:

$$\Rightarrow \begin{bmatrix} (x-1)(x-2) & x(x-1)(x-2) \\ 0 & -(x-1)^2 \end{bmatrix}$$

$$\begin{aligned}
C_2 \rightarrow C_2 - xC_1 &\Rightarrow \begin{bmatrix} (x-1)(x-2) & 0 \\ 0 & -(x-1)^2 \end{bmatrix} \\
R_1 \rightarrow R_1 + R_2 &\Rightarrow \begin{bmatrix} (x-1)(x-2) & (x-1)^2 \\ 0 & -(x-1)^2 \end{bmatrix} \\
C_2 \rightarrow C_2 - C_1 &\Rightarrow \begin{bmatrix} (x-1)(x-2) & x-1 \\ 0 & -(x-1)^2 \end{bmatrix} \\
C_1 \leftrightarrow C_2 &\Rightarrow \begin{bmatrix} x-1 & (x-1)(x-2) \\ -(x-1)^2 & 0 \end{bmatrix} \\
C_2 \rightarrow C_2 - (x-2)C_1 &\Rightarrow \begin{bmatrix} x-1 & 0 \\ -(x-1)^2 & (x-2)(x-1)^2 \end{bmatrix} \\
R_2 \rightarrow R_2 + (x-1)R_1 &\Rightarrow \begin{bmatrix} x-1 & 0 \\ 0 & (x-2)(x-1)^2 \end{bmatrix}
\end{aligned}$$

and here we stop, as we have a matrix in Smith canonical form. Thus

$$xI_4 - B \sim \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & x-1 & \\ & & & (x-1)^2(x-2) \end{bmatrix}$$

so the invariant factors of  $B$  are the non-trivial ones of  $xI_4 - B$ , i.e.

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

Also, the elementary divisors of  $B$  are

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

so the Jordan canonical form of  $B$  is

$$J_2(1) \oplus J_1(1) \oplus J_1(2).$$

### THEOREM 6.9

Let  $A, B \in M_{n \times n}(F)$ . Then  $A$  is similar to  $B$

$$\begin{aligned}
&\Leftrightarrow xI_n - A \text{ is equivalent to } xI_n - B \\
&\Leftrightarrow xI_n - A \text{ and } xI_n - B \text{ have the same} \\
&\quad \text{Smith canonical form.}
\end{aligned}$$

proof

$\Rightarrow$  Obvious. If  $P^{-1}AP = B$ ,  $P \in M_{n \times n}(F)$  then

$$\begin{aligned}P^{-1}(xI_n - A)P &= xI_n - P^{-1}AP \\ &= xI_n - B.\end{aligned}$$

$\Leftarrow$  If  $xI_n - A$  and  $xI_n - B$  are equivalent over  $F[x]$ , then they have the same invariant factors and so have the same non-trivial invariant factors. That is,  $A$  and  $B$  have the same invariant factors and hence are similar.

**Note:** It is possible to start from  $xI_n - A$  and find  $P \in M_{n \times n}(F)$  such that

$$P^{-1}AP = \bigoplus_{k=1}^s C(d_k)$$

where

$$P_1(xI_n - B)Q_1 = \text{diag}(1, \dots, 1, d_1, \dots, d_s).$$

(See Perlis, Theory of matrices, p. 144, Corollary 8-1 and p. 137, Theorem 7-9.)

**THEOREM 6.10**

*Every unit in  $M_{n \times n}(F[x])$  is a product of elementary row and column matrices.*

PROOF: Problem sheet 7, Question 12.