The Smith Normal Form of a Matrix

Patrick J. Morandi

February 17, 2005

In this note we will discuss the structure theorem for finitely generated modules over a principal ideal domain from the point of view of matrices. We will then give a matrix-theoretic proof of the structure theorem from the point of view of the Smith normal form of a matrix over a principal ideal domain. One benefit from this method is that there are algorithms for finding the Smith normal form of a matrix, and these are programmed into common computer algebra packages such as Maple and MuPAD. These packages will make it easy to decompose a finitely generated module over a polynomial ring $F[x]$ into a direct sum of cyclic submodules.

To start, we will need to discuss describing a module by generators and relations. To motivate the definition, let F be a field, and take $A \in M_n(F)$. We can make F^n, viewed as the set of column matrices over F, into an $F[x]$-module by defining $f(x)v = f(A)v$. This module structure is dependent on A; we denote this module by $(F^n)^A$. Write $A = (a_{ij})$. If $\{e_1, \dots, e_n\}$ is the standard basis of F^n, then $xe_j = Ae_j = \sum_{i=1}^n a_{ij}e_i$ for each j. Consequently,

$$\begin{aligned}
(x - a_{11})e_1 - a_{21}e_2 - \cdots - a_{n1}e_n &= \mathbf{0}, \\
-a_{12}e_1 + (x - a_{22})e_2 - \cdots - a_{n2}e_n &= \mathbf{0}, \\
&\vdots \\
-a_{1n}e_1 - \cdots + (x - a_{nn})e_n &= \mathbf{0}.
\end{aligned}$$

The $\{e_i\}$ are generators of $(F^n)^A$ as an $F[x]$-module, and these equations give relations between the generators. Moreover, as we will prove later, the module $(F^n)^A$ is determined by the generators $e_1, \dots, e_n$ and the relations given above.

1 Generators and Relations

Let R be a principal ideal domain and let M be a finitely generated R-module. If $\{m_1, \dots, m_n\}$ is a set of generators of M, then we have a surjective R-module homomorphism $\varphi : R^n \to M$ given by sending $(r_1, \dots, r_n)$ to $\sum_{i=1}^n r_i m_i$. Let K be the kernel of φ. Then $M \cong R^n/K$, a fact we will use repeatedly. If $(r_1, \dots, r_n) \in K$, then $\sum_{i=1}^n r_i m_i = \mathbf{0}$. Thus, an element of K gives rise to a *relation* among the generators $\{m_1, \ldots, m_n\}$. We will refer to K as the *relation submodule* of R^n relative to the generators $m_1, \ldots, m_n$. It is known that K is finitely generated; we will give a proof of this fact for the module $(F^n)^A$ described in the previous section. Suppose that $\{k_1, \ldots, k_m\} \subseteq R^n$ is a generating set for K. If $k_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, then we will refer to the matrix (a_{ij}) over R as the *relation matrix* for M relative to the generating set $\{m_1, \ldots, m_n\}$ of M and the generating set $\{k_1, \ldots, k_m\}$ of K. This matrix has k_i as its i-th row for each i. Since this matrix depends not just on the generating sets for M and K but by the order in which we write the elements, we will use ordered sets, or lists, to denote generating sets. We will write $[m_1, \ldots, m_n]$ to denote an ordered n-tuple.

Generating sets for a module M and for a relation submodule K are not unique. The goal of this section is to see how changing either results in a change in the relation matrix. To get an idea of the general situation, we consider some examples.

Example 1.1. Let $M = \mathbb{Z}_4 \oplus \mathbb{Z}_{12}$. Then M is generated by $m_1 = (1, 0)$ and $m_2 = (0, 1)$. Moreover, $4m_1 = 0$ and $12m_2 = 0$. In fact, if we consider the homomorphism $\varphi : \mathbb{Z}^2 \to M$ sending (r, s) to $rm_1 + sm_2$, then

$$\begin{aligned} \ker(\varphi) &= \left\{(r, s) \in \mathbb{Z}^2 : (r + 4\mathbb{Z}, s + 12\mathbb{Z}) = (0, 0)\right\} \\ &= \{(4a, 12b) : a, b \in \mathbb{Z}\}. \end{aligned}$$

Thus, every element $(4a, 12b)$ in the kernel can be written as $a(4, 0) + b(0, 12)$ for some $a, b \in \mathbb{Z}$. Therefore, $[(4, 0), (0, 12)]$ is an ordered generating set for $\ker(\varphi)$. The relation matrix for this generating set is then the diagonal matrix

$$\begin{pmatrix} 4 & 0 \\ 0 & 12 \end{pmatrix}.$$

Example 1.2. Let the Abelian group M have generators $[m_1, m_2]$, and suppose that the relation submodule K is generated by $[(3, 0), (0, 6)]$. Then the relation matrix is the diagonal matrix

$$\begin{pmatrix} 3 & 0 \\ 0 & 6 \end{pmatrix}.$$

Moreover, the relation submodule K relative to $[m_1, m_2]$ is

$$\begin{aligned} K &= \{a(3, 0) + b(0, 6) : a, b \in \mathbb{Z}\} \\ &= \{(3a, 6b) : a, b \in \mathbb{Z}\}. \end{aligned}$$

Furthermore, K is also the kernel of the map $\sigma : \mathbb{Z}^2 \to \mathbb{Z}_3 \oplus \mathbb{Z}_6$ which is defined by $\sigma(r, s) = (r + 3\mathbb{Z}, s + 6\mathbb{Z})$. Therefore, $\mathbb{Z}^2/K \cong \mathbb{Z}_3 \oplus \mathbb{Z}_6$. However, the meaning of K shows that $M \cong \mathbb{Z}^2/K$. Therefore, $M \cong \mathbb{Z}_3 \oplus \mathbb{Z}_6$. The consequence of this example is that if our relation matrix is diagonal, then we can determine explicitly M as a direct sum of cyclic modules.

Example 1.3. Let the Abelian group M have generators $[m_1, m_2]$, and suppose these generators satisfy the relations $2m_1+4m_2 = 0$ and $-2m_1+6m_2 = 0$. Then the relation submodule K contains $k_1 = (2, 4)$ and $k_2 = (-2, 6)$. If these generate K, the relation matrix is

$$\begin{pmatrix} 2 & 4 \\ -2 & 6 \end{pmatrix}.$$

Note that K is also generated by k_1 and k_1+k_2. These pairs are $(2, 4)$ and $(0, 10)$. Therefore, relative to this new generating set of K, the relation matrix is

$$\begin{pmatrix} 2 & 4 \\ 0 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ -2 & 6 \end{pmatrix}.$$

This new relation matrix is obtained from the original by adding the first row to the second. On the other hand, we can instead use the generating set $[n_1 = m_1 + 2m_2, n_2 = m_2]$. The two relations can be rewritten as $2n_1 = 0$ and $-2n_1 + 10n_2 = 0$. Therefore, with respect to this new generating set, the relation matrix is

$$\begin{pmatrix} 2 & 0 \\ -2 & 10 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ -2 & 6 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

This matrix was obtained from the original by subtracting 2 times the first column from the second column.

The behavior in this example is typical of what happens when we change generators or relations.

Lemma 1.4. *Let M be a finitely generated R-module, with ordered generating set $[m_1, \ldots, m_n]$. Suppose that the relation submodule K is generated by $[k_1, \ldots, k_p]$. Let A be the $p \times n$ relation matrix relative to these generators.*

(1) *Let $P \in M_p(R)$ be an invertible matrix. If $[l_1, \ldots, l_p]$ are the rows of PA, then they generate K, and so PA is the relation matrix relative to $[m_1, \ldots, m_n]$ and $[l_1, \ldots, l_p]$.*

(2) *Let $Q \in M_n(R)$ be an invertible matrix and write $Q^{-1} = (q_{ij})$. If m'_j is defined by $m'_j = \sum_i q_{ij} m_i$ for $1 \le j \le n$, then $[m'_1, \ldots, m'_n]$ is a generating set for M and the rows of AQ generate the corresponding relation submodule. Therefore, AQ is a relation matrix relative to $[m'_1, \ldots, m'_n]$.*

(3) *Let P and Q be $p \times p$ and $n \times n$ invertible matrices, respectively. If $B = PAQ$, then B is the relation matrix relative to an appropriate ordered set of generators of M and of the corresponding relation submodule.*

Proof. (1). The rows of A are the generators $k_1, \ldots, k_p$ of K. If $P = (\alpha_{ij})$, then the rows of PA are

$$\begin{aligned} l_1 &= \alpha_{11}k_1 + \cdots + \alpha_{1p}k_p, \\ l_2 &= \alpha_{21}k_1 + \cdots + \alpha_{2p}k_p, \\ &\vdots \\ l_p &= \alpha_{p1}k_1 + \cdots + \alpha_{pp}kr_p. \end{aligned}$$

The l_i are then elements of K. Moreover, $[l_1, \ldots, l_p]$ is another generating set for K, since we can recover the k_i from the l_j by using P^{-1}: if $P^{-1} = \left(\beta_{ij}\right)$, then $k_i = \beta_{i1}l_1 + \cdots + \beta_{ip}l_p$ for each i. As the rows of PA are then generators for K, this matrix is a relation matrix for M.

(2). The m'_j are generators of M since each of the m_i are linear combinations of the m'_j; in fact, if $Q = (\alpha_{ij})$, then $m_i = \sum_{j=1}^n \alpha_{ij}m'_j$. By thinking about matrix multiplication, the relations for the original generators can be written as a single matrix equation

$$A \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This can be written as

$$(AQ)\, Q^{-1} \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

or

$$AQ \begin{pmatrix} m'_1 \\ \vdots \\ m'_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Therefore, the rows of AQ are relations relative to the new generating set $[m'_1, \ldots, m'_n]$. The rows generate the relation submodule K' relative to the new generating set since if $r = (r_1, \ldots, r_n) \in K'$, then $\sum_{i=1}^n r_i m'_i = 0$. Writing this in terms of matrix multiplication, we have

$$(r_1, \ldots, r_n)Q^{-1} \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

and so the row matrix $(r_1, \ldots, r_n)Q^{-1} \in K$. Thus, $(r_1, \ldots, r_n)Q^{-1} = \sum_{i=1}^p c_i k_i$ for some $c_i \in R$. Multiplying on the right by Q yields $(r_1, \ldots, r_n) = \sum_{i=1}^p c_i\,(k_i Q)$, a linear combination of the rows of AQ. Thus, the rows of AQ do generate the relation submodule.

Finally, (3) simply combines (1) and (2). □

To get some feel for the relevance of this lemma, we recall the connection between row and column operations and matrix multiplication. Consider the three types of row (resp. column) operations:

1. multiplying a row (resp. column) by an invertible element of R;

2. interchanging two rows (resp. columns);

3. adding a multiple of one row (resp. column) to another.

Each of these operations has an inverse operation that undoes the given operation. For example, if we multiply a row by a unit $u \in R$, then we can undo the operation by multiplying the row by u^{-1}. Similarly, if we add α times row i to row j to convert a matrix A to a new matrix B, then we can undo this by adding $-\alpha$ times row i to row j of B to recover A. If E is the matrix obtained by performing a row operation on the $n \times n$ identity matrix, and if A is an $n \times m$ matrix, then EA is the matrix obtained by performing the given row operation on A. Similarly, if E' is the matrix obtained by performing a column operation on the $m \times m$ identity matrix, then AE' is the matrix obtained by performing the given column operation on A. We claim that E and E' are invertible matrices; to see why for E, if G is the matrix obtained by performing the inverse row operation, then $GE = I$, since GEI is the matrix obtained by first performing the row operation on I and then performing the inverse operation. Thus, E is invertible.

As a consequence of this, if we start with a matrix A and perform a series of row and column operations, the resulting matrix will have the form PAQ for some invertible matrices P and Q; the matrix P will be a product of matrices corresponding to to elementary row operations, and Q has a similar description.

Example 1.5. Consider the Abelian group M in the previous example, with generators $[m_1, m_2]$ and relations $2m_1 + 4m_2 = 0$ and $-2m_1 + 6m_2 = 0$. So, relative to the ordered generating sets $[m_1, m_2]$ and $[k_1, k_2] = [(2,4), (-2,6)]$, our relation matrix is

$$\begin{pmatrix} 2 & 4 \\ -2 & 6 \end{pmatrix}.$$

Subtracting 2 times column 1 from column 2 yields the new lists $[m_1, 2m_1 + m_2]$ and $[k_1, k_2]$, with relation matrix

$$\begin{pmatrix} 2 & 0 \\ -2 & 10 \end{pmatrix}.$$

Adding row 1 to row 2 yields

$$\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix},$$

which corresponds to $[m_1, 2m_1 + m_2]$ and $[k_1, k_1 + k_2]$. From this description of M, we see that $M \cong \mathbb{Z}_2 \oplus \mathbb{Z}_{10}$.

We now see that having a diagonal relation matrix allows us to write the module as a direct sum of cyclic modules.

Proposition 1.6. *Suppose that A is a relation matrix for an R-module M. If there are invertible matrices P and Q for which*

$$PAQ = \begin{pmatrix} a_1 & 0 & \cdots & & \\ 0 & a_2 & 0 & \cdots & \\ \vdots & & \ddots & & \\ & & & a_n & \\ 0 & \cdots & & & \end{pmatrix}$$

is a diagonal matrix, then $M \cong R/(a_1) \oplus \cdots \oplus R/(a_n)$.

Proof. The matrix PAQ above is the relation matrix for an ordered generating set $[m_1, \ldots, m_n]$ relative to a relation submodule generated by the rows of PAQ. If $\varphi : R^n \to M$ is the corresponding homomorphism which sends $(r_1, \ldots, r_n)$ to $\sum_{i=1}^n r_i m_i$, then the relation submodule K is the kernel of φ. Thus, $M \cong R^n/K$. However, K is also the kernel of the surjective R-module homomorphism $R^n \to R/(a_1) \oplus \cdots \oplus R/(a_n)$ given by sending $(r_1, \ldots, r_n)$ to $(r_1 + (a_1), \ldots, r_n + (a_n))$. Thus, $R/(a_1) \oplus \cdots \oplus R/(a_n)$ is also isomorphic to R^n/K. Therefore, $M \cong R/(a_1) \oplus \cdots \oplus R/(a_n)$. □

2 The Smith Normal Form

Let R be a principal ideal domain and let A be a $p \times n$ matrix with entries in R. We say that A is in *Smith normal form* if there are nonzero $a_1, \ldots, a_m \in R$ such that a_i divides a_{i+1} for each $i < m$, and for which

$$A = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

We will prove that every matrix over R has a Smith normal form. In the proof we will use a fact about principal ideal domains, stated in Walker: If $(a_1) \subseteq \cdots (a_2) \subseteq \cdots$ is an increasing sequence of ideals, then there is an n such that $(a_n) = (a_{n+1}) = \cdots$. To see why this is true, a short argument proves that the union of the (a_i) is an ideal. Thus, this union is of the form (b) for some b. Now, as $b \in (b)$, we have $b \in \bigcup_{i=1}^\infty (a_i)$. Thus, for some n, we have $b \in (a_n)$. Therefore, as $(a_n) \subseteq (b)$, we get $(a_n) = (b)$. This forces $(a_n) = (a_{n+1}) = \cdots = (b)$.

Theorem 2.1. *If A is a matrix with entries in a principal ideal domain R, then there are invertible matrices P and Q over R such that PAQ is in Smith normal form.*

Proof. To make the proof more clear, we illustrate the idea for 2×2 matrices. Start with a matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Let $e = \gcd(a, c)$, and write $e = ax + cy$ for some $x, y \in R$. Write $a = e\alpha$ and $c = e\beta$ for some $\alpha, \beta \in R$. Then $1 = \alpha x + \beta y$. We have

$$\begin{pmatrix} x & y \\ -\beta & \alpha \end{pmatrix}^{-1} = \begin{pmatrix} \alpha & -y \\ \beta & x \end{pmatrix}.$$

Thus, the matrix

$$\begin{pmatrix} x & 7 \\ -\beta & \alpha \end{pmatrix}$$

is invertible. Moreover,

$$\begin{pmatrix} x & y \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} e & bx + dy \\ -a\beta + c\alpha & -b\beta + d\alpha \end{pmatrix}.$$

Since e divides $-a\beta + c\alpha$, a row operation then reduces this matrix to one of the form

$$\begin{pmatrix} e & u \\ 0 & v \end{pmatrix}.$$

A similar argument, applied to the first row instead of the first column, allows us to multiply on the right by an invertible matrix and obtain a matrix to the form

$$\begin{pmatrix} e_1 & 0 \\ * & * \end{pmatrix},$$

where $e_1 = \gcd(e, u)$. Continuing this process, alternating between the first row and the first column, will produce a sequence of elements $e, e_1, \ldots$ such that e_1 divides e, e_2 divides e_1, and so on. In terms of ideals, this says $(e) \subseteq (e_1) \subseteq \cdots$. Because any increasing sequence of principal ideals stabilizes in a principal ideal domain, we must arrive, in finitely many steps, with a matrix of the form

$$\begin{pmatrix} f & 0 \\ g & h \end{pmatrix} \text{ or } \begin{pmatrix} f & g \\ 0 & h \end{pmatrix}$$

in which f divides g. One more row or column operation will then yield a matrix of the form

$$\begin{pmatrix} f & 0 \\ 0 & k \end{pmatrix}.$$

Thus, by multiplying on the left and right by invertible matrices, we obtain a diagonal matrix.

Once we have reduced to a diagonal matrix

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

to get the Smith normal form, let $d = \gcd(a, b)$. We may write $d = ax + by$ for some $x, y \in R$. Moreover, write $a = d\alpha$ and $b = d\beta$ for some $\alpha, \beta \in R$. We then perform the following row and column operations, yielding

$$\begin{aligned} \begin{pmatrix} a\, 0 \\ 0\, b \end{pmatrix} \longrightarrow \begin{pmatrix} a & 0 \\ ax & b \end{pmatrix} \longrightarrow \begin{pmatrix} a & 0 \\ ax + by & b \end{pmatrix} = \begin{pmatrix} a\, 0 \\ d\, b \end{pmatrix} \\ \longrightarrow \begin{pmatrix} 0 & -b\alpha \\ d & b \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & -b\alpha \\ d & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} d & 0 \\ 0 & -b\alpha \end{pmatrix}, \end{aligned}$$

a diagonal matrix in Smith normal form since d divides $-b\alpha$. □

As a consequence of the existence of a Smith normal form, we obtain the structure theorem for finitely generated modules over a principal ideal domain.

Corollary 2.2. *If M is a finitely generated module over a principal ideal domain R, then there are elements $a_1, \ldots, a_m \in R$ such that a_i divides a_{i+1} for each $i = 1, \ldots, m-1$, and an integer $t \geq 0$ such that $M \cong R/(a_1) \oplus \cdots \oplus R/(a_m) \oplus R^t$.*

Proof. Let A be a relation matrix for M, and let B be its Smith normal form. Then $B = PAQ$ for some invertible matrices P, Q. If

$$B = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

Proposition 1.6 then shows that

$$\begin{aligned} M &\cong R/(a_1) \oplus \cdots \oplus R/(a_m) \oplus R/(0) \oplus \cdots R/(0) \\ &\cong R/(a_1) \oplus \cdots \oplus R/(a_m) \oplus R^t \end{aligned}$$

for some $t \geq 0$. □

Example 2.3. Let A be the Abelian group with generators m_1, m_2, m_3 with relation submodule generated by $(8, 4, 8), (4, 8, 4)$. Then the basic relations are

$$8m_1 + 4m_2 + 8m_3 = 0,$$
$$4m_1 + 8m_2 + 4m_3 = 0.$$

The corresponding relation matrix is

$$\begin{pmatrix} 8 & 4 & 8 \\ 4 & 8 & 4 \end{pmatrix}.$$

By performing row and column operations, we reduce this matrix to Smith normal form and list the effect on the generators of the group and the corresponding relation subgroup.

matrix	**generators**	**relations**
$\begin{pmatrix} 8 & 4 & 8 \\ 4 & 8 & 4 \end{pmatrix}$	m_1, m_2, m_3	$8m_1 + 4m_2 + 8m_3 = 0,$ $4m_1 + 8m_2 + 4m_3 = 0.$
$\begin{pmatrix} 0 & -12 & 0 \\ 4 & 8 & 4 \end{pmatrix}$	m_1, m_2, m_3	$-12m_2 = 0,$ $4m_1 + 8m_2 + 4m_3 = 0.$
$\begin{pmatrix} 0 & -12 & 0 \\ 4 & 0 & 4 \end{pmatrix}$	$m_1 + 2m_2, m_2, m_3$	$-12m_2 = 0$ $4(m_1 + 2m_2) + 4m_3 = 0$
$\begin{pmatrix} 0 & -12 & 0 \\ 4 & 0 & 0 \end{pmatrix}$	$m_1 + 2m_2 + m_3, m_2$	$-12m_2 = 0$ $4(m_1 + 2m_2 + m_3) = 0$
$\begin{pmatrix} 4 & 0 & 0 \\ 0 & -12 & 0 \end{pmatrix}$	$m_2, m_1 + 2m_2 + m_3$	$4(m_1 + 2m_2 + m_3) = 0$ $-12m_2 = 0$
$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 12 & 0 \end{pmatrix}$	$-m_2, m_1 + 2m_2 + m_3$	$4(m_1 + 2m_2 + m_3) = 0$ $12(-m_2) = 0$

From the final matrix, we see that $A \cong \mathbb{Z}_4 \oplus \mathbb{Z}_{12}$.

We now specialize to the case of modules over the polynomial ring $F[x]$ over a field F. Let $A \in M_n(F)$ be a matrix, and consider the module $(F^n)^A$ by making F^n into an $F[x]$-module via the scalar multiplication $f(x) \cdot m = f(A)m$. Then $(F^n)^A$ is a finitely generated module over the principal ideal domain $F[x]$. Let $e_1, \ldots, e_n$ be the standard basis for F^n. Consider the $F[x]$-module homomorphism $\varphi : F[x]^n \to F^n$ which sends $(f_1(x), \ldots, f_n(x))$ to $\sum_{i=1}^n f_i(x)e_i$. We wish to determine generators for $\ker(\varphi)$ in order to apply the results of the previous section. Referring to the beginning of the note, if $A = (a_{ij})$, then the generators e_i satisfy the relations

$$\begin{aligned} (x - a_{11})e_1 - a_{21}e_2 - \cdots - a_{n1}e_n &= \mathbf{0}, \\ -a_{12}e_1 + (x - a_{22})e_2 - \cdots - a_{n2}e_n &= \mathbf{0}, \\ &\vdots \\ -a_{1n}e_1 - \cdots + (x - a_{nn})e_n &= \mathbf{0}. \end{aligned}$$

Building a matrix from the coefficients yields

$$\begin{pmatrix} x - a_{11} & -a_{21} & \cdots & -a_{n1} \\ -a_{12} & x - a_{22} & \cdots & -a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & x - a_{nn} \end{pmatrix} = xI - A^T.$$

Thus, the rows of $xI - A^T$ are elements of the relation submodule of $(F^n)^A$ relative to $[e_1, \ldots, e_n]$. We will prove that $xI - A^T$ is a relation matrix for $(F^n)^A$ relative to the generating set $[e_1, \ldots, e_n]$. This amounts to proving that the rows of $xI - A^T$ generates the relation submodule. Thus, finding the Smith normal form of $xI - A^T$ will show how to write $(F^n)^A$ as a direct sum of cyclic modules.

Let $v_1, \ldots, v_n$ be the rows of $xI - A^T$, and let $E_1, \ldots, E_n$ be the standard basis vectors of $F[x]^n$.

Lemma 2.4. *Let $\sum_{i=1}^n f_i(x)E_i \in F[x]^n$. Then there are $g_i(x) \in F[x]$ and $\alpha_i \in F$ such that $\sum_{i=1}^n f_i(x)E_i = \sum_{i=1}^n g_i(x)v_i + \sum_{i=1}^n \alpha_i E_i$.*

Proof. We prove this by inducting on the maximum m of the degrees of the $f_i(x)$. The case $m = 0$ is trivial, since in this case each $f_i(x)$ is a constant polynomial, and then we can choose $g_i(x) = 0$ and $\alpha_i = f_i(x) \in F$. Next, suppose that $m > 0$ and that the result holds for vectors of polynomials whose maximum degree is $< m$. By the division algorithm, we may write $f_1(x) = q_1(x)(x - a_{11}) + r_1$ for some $q_1(x) \in F[x]$ and $r_1 \in F$. Then

$$\begin{aligned}(f_1(x), 0, \ldots, 0) &= (q_1(x)(x - a_{11}) + r_1, 0, \ldots, 0) \\ &= q_1(x)\,(x - a_{11}, -a_{21}, \ldots, -a_{n1}) + (r_1, q_1(x)a_{21}, \ldots, q_1(x)a_{n1}) \\ &= q_1(x)v_1 + (r_1, q_1(x)a_{21}, \ldots, q_1(x)a_{n1}).\end{aligned}$$

Note that $\deg(q_1(x)) = \deg(f_1(x)) - 1$. Therefore, each entry of the second vector has degree strictly less than $\deg(f_1(x))$. Repeating this idea for each $f_i(x)$ and subsequently rewriting each $f_i(x)E_i$, we see that

$$\sum_{i=1}^n f_i(x)E_i = \sum_{i=1}^n q_i(x)v_i + \sum_{i=1}^n h_i(x)E_i$$

for some $h_i(x) \in F[x]$ with $\deg(h_i(x)) < M$. By induction, we may write $\sum_{i=1}^n h_i(x)E_i = \sum_{i=1}^n k_i(x)v_i + \sum_{i=1}^n \alpha_i E_i$ for some $k_i(x) \in F[x]$ and $\alpha_i \in F$. Then

$$\sum_{i=1}^n f_i(x)E_i = \sum_{i=1}^n (q_i(x) + k_i(x)v_i + \sum_{i=1}^n \alpha_i E_i,$$

which is of the desired form. Thus, the lemma follows by induction. □

Proposition 2.5. *If $\varphi : F[x]^n \to (F^n)^A$ is the $F[x]$-module homomorphism defined by $\varphi(f_1(x), \ldots, f_n(x)) = \sum_{i=1}^n f_i(x)e_i$, then the kernel of φ is generated by the rows of $xI - A^T$.*

Proof. To determine the kernel of φ, let L be the submodule of $F[x]^n$ generated by the rows $v_1, \ldots, v_n$ of $xI - A$. We have noted that each $v_i \in \ker(\varphi)$; thus, $L \subseteq \ker(\varphi)$. For the reverse inclusion, suppose that $\sum_{i=1}^n f_i(x)E_i \in \ker(\varphi)$. By the lemma, we may write $\sum_{i=1}^n f_i(x)E_i = \sum_{i=1}^n g_i(x)v_i + \sum_{i=1}^n \alpha_i E_i$ for some $g_i(x) \in F[x]$ and $\alpha_i \in F$. Since each $v_i \in \ker(\varphi)$, we conclude that $\sum_{i=1}^n \alpha_i E_i \in \ker(\varphi)$. However, this element maps to $(\alpha_1, \ldots, \alpha_n) \in F^n$. Consequently, each $\alpha_i = 0$. Therefore, $\sum_{i=1}^n f_i(x)E_i = \sum_{i=1}^n g_i(x)v_i \in L$. □

Corollary 2.6. *Let* $A \in M_n(F)$*, and let* $(F^n)^A$ *be the* $F[x]$*-module via the matrix* A*, as above. If*

$$B = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & f_1(x) & & \\ & & & & \ddots & \\ & & & & & f_m(x) \end{pmatrix}$$

is the Smith normal form of A*, then the* $(F^n)^A \cong F[x]/(f_1(x)) \oplus \cdots \oplus F[x]/(f_m(x))$ *as* $F[x]$*-modules. Thus, the invariant factors of* F^n *are* $f_1(x), \ldots, f_m(x)$*.*

Proof. If B has the form above, then as $F[x]/(1)$ is the zero module, we get the desired decomposition of $(F^n)^A$. □