THE JORDAN CANONICAL FORM OF A KRONECKER PRODUCT

by

PAKO RAMASU

B. Ed (Sc) (Mathematics) University of Botswana, 2006

A THESIS SUBMITTED IN PARTIAL FULFFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the Graduate Academic Unit of Mathematics and Statistics

Supervisor: Dr. Colin Ingalls, PhD, Math & Stats

Examining Board: Dr. Barry Monson, PhD, Mathematics

This thesis is accepted by the

Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

April 2009

ABSTRACT

In 1934, Aitken and Roth studied the Jordan canonical form of a Kronecker product. In this thesis, we use their method to construct the Jordan canonical form of a Kronecker product of two matrices whose eigenvalues are not necessarily distinct. We also use combinatorics and graph theory which were presented by Brualdi in 1985 to derive determinantal divisors of the Kronecker product of two matrices. From this we obtain the elementary divisors and the Jordan canonical form.

Table of Contents

Chapter 1: Introduction

In this chapter, we review basic background information that is needed for the evaluation of the Jordan canonical form of the Kronecker product matrix. We give key definitions and describe standard operations on matrices. We survey some key theorems and propositions, together with their proofs and give some examples. Part of this chapter discusses the matrix $A \otimes B$, the Kronecker product of matrices A and B. We then conclude this chapter by discussing how to construct the Jordan canonical structure of a matrix A. It has been shown that for a matrix A with distinct eigenvalues, this matrix is diagonalizable. If a matrix A has repeated eigenvalues, then its Jordan structure may contain Jordan blocks, this matrix is quasi-diagonal in the sense that its only non-zero entries lie on the diagonal and the superdiagonal.

1.1 Standard Operations

In this section we review some important aspects of the theory of matrices over a field **K**. Most often, **K** will be the real field **R** or complex field **C**. Typically matrices are denoted by upper case letters A, B, etc. The vector space $\mathbf{K}^n$ consists of all n x 1 column vectors over **K**, which will be denoted by lower case bold letters $\boldsymbol{v}$, $\boldsymbol{w}$ etc.

Definition 1.1.1:

The *determinant* of an n x n matrix A = $[a_{ij}]$ is given for any fixed i by

$$\det(A) = |A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$$

where M_{ij} denotes the (n-1) x (n-1) submatrix of the matrix A obtained by deleting the i^{th} row and j^{th} column of A. The expression $(-1)^{i+j}\det(M_{ij})$ is called the *cofactor* of the element a_{ij}. We note that this expansion is independent of the row index i

[Pool, p. 264].

Note that if a matrix A has a zero row, then by cofactor expansion along that row, we get

$$\det(A) = 0.$$

Definition 1.1.2:

For a 2 x 2 matrix A, the determinant is given, in various notations, by

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Definition 1.1.3:

Let A be an n x n matrix. A scalar γ is an *eigenvalue* of a matrix A, if there is a nonzero column vector $\boldsymbol{v} \in \mathbf{K}^n$, such that

$$A\boldsymbol{v} = \gamma\boldsymbol{v}. \tag{1}$$

The vector $\boldsymbol{v}$ is called an *eigenvector* of a matrix A corresponding to the eigenvalue γ.

From equation (1) we have

$$A\boldsymbol{v} = \gamma\boldsymbol{v}$$

$$A\boldsymbol{v} - \gamma\boldsymbol{v} = \mathbf{0}$$

$$(A - \gamma I_n)\boldsymbol{v} = \mathbf{0}. \tag{2}$$

Equation (2) has non-trivial solutions if and only if

$$\det(A - \gamma I_n) = 0. \tag{3}$$

Note that, if $\boldsymbol{v}$ is an eigenvector and δ is a nonzero scalar, then $\delta\boldsymbol{v}$ is also an eigenvector. It is important again to stress out that an eigenvector should be non-zero, since (1) is trivially satisfied by the zero vector for any number γ.

Definition 1.1.4:

Referring to equation (3) above,

$$p(x) = \det(A - xI_n)$$

is a polynomial of degree n called the *characteristic polynomial* of the n x n matrix A. The *characteristic equation* is given by

$$p(x) = 0.$$

Definition 1.1.5:

If

$$q(x) = b_0 + b_1 x + b_2 x^2 + \text{- - -} + b_r x^r$$

is any polynomial in $\mathbf{K}[x]$, and A is any n x n matrix over **K**, then

$$q(A) = b_0 I_n + b_1 A + b_2 A^2 + \text{- - -} + b_r A^r.$$

The *minimal polynomial* of a matrix A is the monic polynomial $t(x)$ of smallest positive degree m, such that

$$t(A) = 0.$$

(Here 0 denotes the n x n zero matrix).

Proposition 1.1.6:

The minimal polynomial $t(x)$ exists and is unique. Moreover, the minimal polynomial divides every polynomial $q(x)$ for which

$$q(A) = 0$$

[Laub, p. 76].

Cayley – Hamilton Theorem 1.1.7:

Let A be an n x n matrix with characteristic polynomial

$$p(x) = x^n + c_{n-1}x^{n-1} + \text{- - -} + c_1 x + c_0.$$

Then

$$p(A) = A^n + c_{n-1}A^{n-1} + \text{- - -} + c_1 A + c_0 I_n = 0.$$

Proof

Let B be the adjoint of the matrix A, namely the transpose of the matrix of cofactors.

Then from Definition 1.1.1,

$$B = [b_{ij}] = [(-1)^{i+j}\det(M_{ij})].$$

Recall that

$$AB = \det(A)I_n. \tag{4}$$

This result is true for any square matrix with entries in any commutative ring and therefore it is true in particular for the matrix $(xI_n - A)$. The adjoint $B(x)$ of $(xI_n - A)$ is an x-matrix, each element which is a polynomial of degree (n-1) or less in x, since each such entry is a minor of order $(n-1)$ for $(xI_n - A)$. In general

$$B(x) = B_{n-1}x^{n-1} + B_{n-2}x^{n-2} + \text{- - -} + B_1x + B_0 \tag{5}$$

where the B_i are n x n matrices with constant entries. Applying the result (4) to $(xI_n - A)$ we have

$$\begin{aligned}\det(xI_n - A)I_n &= (xI_n - A)(B_{n-1}x^{n-1} + B_{n-2}x^{n-2} + \text{- - -} + B_1x + B_0)\\ &= (x^n + c_{n-1}x^{n-1} + \text{- - -} + c_0)I_n. \end{aligned} \tag{6}$$

Equating coefficients of powers of x we get

$$\begin{aligned} B_{n-1} &= I_n \\ B_{n-2} - AB_{n-1} &= c_{n-1}I_n \\ &\vdots \\ -AB_0 &= c_0 I_n \end{aligned}$$

Multiplying these equations by A^n, A^{n-1}, - - -,I_n and summing gives

$$A^n + c_{n-1}A^{n-1} + \text{- - -} + c_0 I_n = 0 \ \blacksquare$$

[Wilk, p. 38-39].

We conclude that every square matrix satisfies its own characteristic equation. Therefore, the minimal polynomial of a matrix cannot be of degree greater than that of the characteristic polynomial, and the minimal polynomial divides the characteristic polynomial p(x).

Later on, we will show that the Jordan canonical form of a matrix A determines the minimal polynomial of a matrix A, but the converse is not true.

<u>Definition 1.1.8:</u>

For each eigenvalue γ of an n x n matrix A, we have the eigenvector $\boldsymbol{v}$ such that

$$(A - \gamma I_n)\boldsymbol{v} = \mathbf{0}.$$

The kernel of the matrix $A - \gamma I_n$ contains each such vector $\boldsymbol{v}$ and is called the *eigenspace* associated with the eigenvalue γ, denoted by Eγ:

$$E\gamma = \ker(A - \gamma I_n).$$

Furthermore, the dimension of the eigenspace Eγ is called the *geometric multiplicity* corresponding to the eigenvalue γ. In other words, the geometric multiplicity of the eigenvalue γ is the nullity of the matrix $A - \gamma I_n$.

<u>Definition 1.1.9:</u>

Let I_n be the n x n identity matrix and let A be an n x n matrix. A matrix B of order n is called the *inverse matrix* of a matrix A if and only if

$$AB = I_n = BA.$$

Since this condition uniquely determines A, we just write

$$B = A^{-1}.$$

Furthermore, if matrix A^{-1} exists, then

$$A\,A^{-1} = I_n,$$

so that

$$\det(A) \neq 0.$$

We shall say that a matrix A is *non-singular* or *invertible* if a matrix A^{-1} exists; otherwise, the matrix A is *singular*.

Example 1.1.10:

Find the eigenvalues and the corresponding eigenspaces for the following matrix

$$A = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}.$$

Solution

We first simplify the characteristic polynomial

$$\det(A - \gamma I) = -(10 - \gamma)(11 + \gamma) + 108$$

$$= -(110 + 10\gamma - 11\gamma - \gamma^2) + 108$$

$$= \gamma^2 + \gamma - 2$$

$$= (\gamma - 1)(\gamma + 2).$$

The characteristic equation is

$$(\gamma - 1)(\gamma + 2) = 0.$$

It is an easy exercise to confirm the Cayley Hamilton theorem by verifying that

$$p(A) = A^2 + A - 2I$$

$$= 0.$$

Since the characteristic polynomial of A has no repeated factors, the minimal polynomial of A is equal the characteristic polynomial.

We have the eigenvalues

$$\gamma_1 = 1, \quad \gamma_2 = -2.$$

To find an eigenspace for $\gamma_1 = 1$, we solve for the eigenvector

$$(A - \gamma_1 I_2)\boldsymbol{v}_1 = \mathbf{0}.$$

We obtain the redundant system

$$9x - 18y = 0$$

$$6x - 12y = 0$$

Thus a non-trivial eigenvector for γ_1 is

$$\boldsymbol{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

For γ_2, we similarly get

$$\boldsymbol{v}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

In the next section, we describe a non-singular matrix P such that

$$P^{-1}AP = J_A$$

is the Jordan canonical form for A. Here we make

$$P = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix},$$

whose columns are the above eigenvectors. It is then easy to check that

$$P^{-1}AP = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

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

$$= J_A.$$

Thus A is diagonalizable.

Definition 1.1.11:

A matrix A is in *row echelon* form if it satisfies the following conditions

(a) If a row is not entirely made out of 0's, then the first non-zero number in the row must equal 1. (This entry is called a leading 1).

(b) If there are any rows entirely made out of 0's, then they are grouped at the bottom of the matrix.

(c) In any two successive rows that do not consist of 0's, the leading 1 in the lower row occurs further to the right than the leading 1 in the higher row.

Definition 1.1.12:

The *rank* of the m x n matrix A is the dimension of its column space, that is, the dimension of the subspace of $\mathbf{K}^n$ spanned by the columns of A. One can show that this equals the dimension of the space spanned by the rows of A. It is easy to see that this in turn equals the number of non-zero rows in its row echelon form after using elementary row operators to change the matrix A into row echelon form [Pool, p. 75].

Theorem 1.1.13:

If A is an m x n matrix, then

$$\mathrm{rank}(A) + \dim(\ker(A)) = n.$$

Proof

Suppose A has rank w and let R be the reduced row echelon form of the matrix A. Then R has w leading 1's. Now solve the homogeneous system defined by

$$A\mathbf{x} = \mathbf{0}.$$

Then there are w leading variables and n-w free variables in the standard description of the solution. Hence

$$\dim(\ker(A)) = n - w.$$

This gives us

$$\begin{aligned}\mathrm{rank}(A) + \dim(\ker(A)) &= w + (n - w)\\ &= n. \blacksquare\end{aligned}$$

[Pool, p. 203, Th. 3.26].

Definition 1.1.14:

The *algebraic multiplicity* of an eigenvalue γ for the matrix A is the multiplicity of γ as a root of the characteristic equation $p(x) = 0$.

Recall that the *geometric multiplicity* of an eigenvalue γ is the number of linearly independent eigenvectors associated with γ, that is, the dimension of the eigenspace $E\gamma$.

It is known that, if γ is an eigenvalue of a matrix A of algebraic multiplicity z and geometric multiplicity y, then we must have

$$1 \leq y \leq z$$

[Laub, p. 76].

I will show in Chapter two that the geometric multiplicity of an eigenvalue γ gives the number of Jordan blocks associated with γ, and for a matrix which has a diagonal form then the algebraic multiplicity and geometric multiplicity are the same for each eigenvalue.

<u>Definition 1.1.15:</u>

An n x n matrix A is said to be *defective* if it has an eigenvalue whose geometric multiplicity is less than its algebraic multiplicity. Equivalently, a matrix A is *defective* if it does not have n linearly independent eigenvectors [Laub, p.76].

1.2 The Kronecker Product of Matrices

This section provides an introduction to the Kronecker product (tensor product) of two matrices of arbitrary sizes over a field **K**.

Definition 1.2.1:

Let A = [a_{ij}] be n x n matrix and B = [b_{rs}] be p x q matrix. The *Kronecker product matrix* A$\otimes$B of the matrices A and B is the np x nq matrix given in block format by

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix}$$

[Laub, p. 140].

Example 1.2.2:

Suppose $\boldsymbol{x} \in \mathbf{K}^n$ and $\boldsymbol{y} \in \mathbf{K}^m$ are column vectors. In other words,

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

are n x 1 and m x 1 matrices, respectively. Hence their Kronecker product is the mn x 1 matrix

$$\boldsymbol{x} \otimes \boldsymbol{y} = \begin{bmatrix} x_1\boldsymbol{y} \\ \vdots \\ x_n\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} x_1y_1 \\ x_1y_2 \\ \vdots \\ x_ny_{m-1} \\ x_ny_m \end{bmatrix},$$

which we may take to be a vector in $\mathbf{K}^{nm}$.

Proposition 1.2.3:

In the following it is assumed that A, B, C and D are matrices over a field **K**, occasionally with dimensions chosen to ensure certain operations are defined. Then:

1. The Kronecker product is a bilinear operator. Given $\delta \in$**K**

$$A \otimes (\delta B) = \delta(A \otimes B) = (\delta A) \otimes B$$

$$(A + B) \otimes C = (A \otimes C + (B \otimes C)$$

$$A \otimes (B + C) = \left(A \otimes B\right) + (A \otimes C).$$

2. The Kronecker product is associative, that is

$$(A \otimes B) \otimes C = A \otimes (B \otimes C).$$

3. The Kronecker product is not always commutative, that is, usually

$$A \otimes B \neq B \otimes A.$$

4. Transpose distributes over the Kronecker product:

$$(A \otimes B)^{\mathrm{T}} = A^{\mathrm{T}} \otimes B^{\mathrm{T}}.$$

5. Let A be an m x n matrix, B an r x s matrix, C an n x p matrix and D an s x t matrix. Then

$$\left(A \otimes B\right)(C \otimes D = AC \otimes BD.$$

Note that, $AC \otimes BD$ is an mr x pt matrix.

6. Let I_n and I_m be identity matrices of order n and m respectively; then

$$I_n \otimes I_m = I_{nm}$$

where I_{nm} is identity matrix of order nm.

7. If A and B are square invertible matrices, then

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

8. The determinant of the Kronecker product of an n x n matrix A and an m x m matrix B is given by

$$\det(A \otimes B) = \det(A)^{m} \cdot \det(B)^{n}.$$

9. The trace of the Kronecker product is given by

$$\operatorname{trace}(A \otimes B) = \operatorname{trace}(A) \times \operatorname{trace}(B).$$

10. $\operatorname{rank}(A \otimes B) = (\operatorname{rank}(A))(\operatorname{rank}(B))$.

11. Let $A_1, A_2, ---, A_p$ and $B_1, B_2, ---, B_q$ be given square matrices.

Then

$$(A_1 \oplus A_2 \oplus \ldots \oplus A_p) \otimes (B_1 \oplus B_2 \oplus \ldots \oplus B_q) = \bigoplus_{i=1}^{p} \bigoplus_{j=1}^{q} A_i \otimes B_j. \qquad (1)$$

Note that if $(A_1 \oplus A_2 \oplus \ldots \oplus A_p)$ is an m x m matrix and $(B_1 \oplus B_2 \oplus \ldots \oplus B_q)$ is an n x n matrix then the order of the matrix in (1) is mn.

Proof

These are mainly routine calculations. For part (5) we note

$$(A \otimes B)(C \otimes D = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} c_{11}D & \cdots & c_{1p}D \\ \vdots & \ddots & \vdots \\ c_{n1}D & \cdots & c_{np}D \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{n} a_{1k} c_{k1} BD & \cdots & \sum_{k=1}^{n} a_{1k} c_{kp} BD \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{mk} c_{kl} BD & \cdots & \sum_{k=1}^{n} a_{mk} c_{kp} BD \end{bmatrix}$$

$$= (AC \otimes BD)$$

For part (7), we use part (5) and (6) to get

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = I \otimes I = I \ \blacksquare$$

[Laub, p. 140].

Theorem 1.2.4:

Let A be an n x n matrix with an eigenvalue γ and B an m x m matrix with an eigenvalue μ. Then $\gamma\mu$ is an eigenvalue for $A \otimes B$.

Moreover, if $\boldsymbol{x}_1, \boldsymbol{x}_2$, - - - , $\boldsymbol{x}_p$ are linearly independent eigenvectors for A and $\boldsymbol{y}_1, \boldsymbol{y}_2$, - - - , $\boldsymbol{y}_q$ are linearly independent eigenvectors for B, then

$$\boldsymbol{x}_i \otimes \boldsymbol{y}_j, \ 1 \leq i \leq p \text{ and } 1 \leq j \leq q,$$

are linearly independent eigenvectors for $A \otimes B$.

If A is diagonalizable with eigenvalues γ_1, γ_2, - - -, γ_n and B is diagonalizable with eigenvalues μ_1, μ_2, - - -, μ_m, then the products

$$\gamma_i \mu_j, \ 1 \leq i \leq \text{n and } 1 \leq j \leq \text{m},$$

give all eigenvalues for $A \otimes B$.

Proof

Suppose

$$A\boldsymbol{x} = \gamma\boldsymbol{x},\ \ B\boldsymbol{y} = \mu\boldsymbol{y}, \quad \text{for } \boldsymbol{x} \in \mathbf{K}^{n} \text{ and } \boldsymbol{y} \in \mathbf{K}^{m}.$$

Then by Proposition 1.2.3, part (5),

$$\begin{aligned}(A \otimes B)(\boldsymbol{x} \otimes \boldsymbol{y}) &= A\boldsymbol{x} \otimes B\boldsymbol{y} \\ &= \gamma\boldsymbol{x} \otimes \mu\boldsymbol{y} \\ &= \gamma\mu(\boldsymbol{x} \otimes \boldsymbol{y}).\end{aligned}$$

This proves our first claim. Next suppose $c_{ij} \in \mathbf{K}$ with

$$\sum_{j=1}^{q}\sum_{i=1}^{p} c_{ij}\,(\boldsymbol{x}_i \otimes \boldsymbol{y}_j) = \mathbf{0} \in \mathbf{K}^{nm}.$$

Letting

$$\boldsymbol{w}_j = \sum_{i=1}^{p} c_{ij}\,\boldsymbol{x}_i = \begin{bmatrix} w_{1j} \\ \vdots \\ w_{nj} \end{bmatrix}, \tag{2}$$

say, we obtain

$$\sum_{j=1}^{q} \boldsymbol{w}_j \otimes \boldsymbol{y}_j = \mathbf{0}, \text{ or}$$

$$\begin{bmatrix} \vdots \\ \sum_{j=1}^{q} w_{ij}\mathbf{y}_j \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}.$$

Thus

$$\sum_{j=1}^{q} w_{ij}\,\boldsymbol{y}_j = \mathbf{0}.$$

Since the $\boldsymbol{y}_j$ are independent, all

$$w_{ij} = 0.$$

From (2) the independence of the $\boldsymbol{x}_i$ gives all

$$c_{ij} = 0.$$

Hence the vectors $\boldsymbol{x}_i \otimes \boldsymbol{y}_j$ are independent. Our final claim now follows, since a matrix is diagonalizable if and only if it admits a basis of eigenvectors (see Proposition 1.3.1). ■

In fact, it is true for any matrices A and B over the complex field **C**, that all eigenvalues for A ⊗ B have the form $\gamma_i \mu_j$, as suggested in the Theorem. However, to prove this we must employ the Jordan canonical forms of A and B, as described in the next Chapter.

<u>Example 1.2.5:</u>

$$\mathrm{A} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}, \quad \mathrm{B} = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix}.$$

Then the Kronecker product is given by

$$\mathrm{A} \otimes \mathrm{B} = \begin{bmatrix} 5 & -10 & 4 & -8 \\ 5 & 20 & 4 & 16 \\ 1 & -2 & 2 & -4 \\ 1 & 4 & 2 & 8 \end{bmatrix}.$$

The matrix A has the eigenvalues $\gamma_1 = 6$ and $\gamma_2 = 1$, with eigenvectors

$$\boldsymbol{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \text{ and } \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

respectively.

The matrix B has eigenvalues $\mu_1 = 3$ and $u_2 = 2$, with eigenvectors

$$\boldsymbol{w}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ and } \boldsymbol{w}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

respectively.

Then according to the above Theorem we obtain the eigenvalues and their corresponding eigenvectors of the Kronecker product matrix A $\otimes$ B as follows:

<u>The first eigenvalue is given by</u>

$$\beta_1 = \gamma_1 \mu_1 = 6 \times 3 = 18,$$

with eigenvector

$$\boldsymbol{x}_1 = \boldsymbol{v}_1 \otimes \boldsymbol{w}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$= [-4 \quad 4 \quad -1 \quad 1]^{\mathrm{T}}$$

<u>The second eigenvalue is given by</u>

$$\beta_2 = \gamma_1 \mu_2 = 6 \times 2 = 12$$

with eigenvector

$$\boldsymbol{x}_2 = \boldsymbol{v}_1 \otimes \boldsymbol{w}_2 = [-8 \quad 4 \quad -2 \quad 1]^{\mathrm{T}}.$$

<u>The third eigenvalue</u>

$$\beta_3 = 3$$

with eigenvector

$$\boldsymbol{x}_3 = [1 \quad -1 \quad -1 \quad 1]^{\mathrm{T}}$$

<u>The last eigenvalue is given by</u>

$$\beta_4 = 2$$

with eigenvector

$$\boldsymbol{x}_4 = [2 \quad -1 \quad -2 \quad 1]^{\mathrm{T}}.$$

In the next section, we discuss how a non-singular matrix composed of the eigenvectors corresponding to the eigenvalues of a diagonalizable matrix A gives the Jordan canonical form of a matrix A.

As shown by the above example, the Kronecker product is very useful in generating large and important matrices. In particular, if we know the eigensystems of the matrices A and B, we can easily compute the eigenvectors and the eigenvalues of the Kronecker products $A \otimes B$, $(A \otimes B) \otimes (A \otimes B)$ and so on. It is this property that makes it simple to find the diagonal form of the Kronecker product $A \otimes B$, and we will discuss this in details in the next sections.

1.3 The Jordan Canonical Form

In this section we look at the Jordan canonical form of a matrix, introduced by Jordan in 1870 with the aim of simplifying the discussion of linear substitutions. Most of the material covered in this section is paraphrased from Ortega [Orte, p. 117-129].

Proposition 1.3.1:

Let A be an n x n matrix. A is similar to the diagonal matrix given by

$$J_A := \mathrm{diag}(\gamma_1, \gamma_2, ---, \gamma_n) \tag{1}$$

if and only if A has n linearly independent eigenvectors $\boldsymbol{x}_1, \boldsymbol{x}_2, - - - , \boldsymbol{x}_n$.

If γ_i, i = 1, 2, - - - , n, are the corresponding eigenvalues, then

$$S^{-1}AS = J_A = \mathrm{diag}(\gamma_1, \gamma_2, - - - , \gamma_n),$$

where

$$S = [\boldsymbol{x}_1, \boldsymbol{x}_2, - - - , \boldsymbol{x}_n]$$

is the matrix whose columns are the eigenvectors.

Proof

First we compute

$$\begin{aligned} AS &= A[\boldsymbol{x}_1, \boldsymbol{x}_2, - - - , \boldsymbol{x}_n] \\ &= [A\boldsymbol{x}_1, A\boldsymbol{x}_2, - - - , A\boldsymbol{x}_n] \\ &= [\gamma_1\boldsymbol{x}_1, \gamma_2\boldsymbol{x}_2, - - - , \gamma_n\boldsymbol{x}_n] \\ &= [\boldsymbol{x}_1, \boldsymbol{x}_2, - - - , \boldsymbol{x}_n]J_A \end{aligned}$$

$$= SJ_A.$$

Then the diagonalization is complete, and we have

$$S^{-1}AS = J_A.$$

Example 1.3.2:

Determine the Jordan canonical form of the matrix

$$A = \begin{bmatrix} 3 & 5 \\ -2 & -4 \end{bmatrix}.$$

Solution

The matrix A has the eigenvalues $\gamma_1 = 1$ and $\gamma_2 = -2$. We find eigenvectors

$$\boldsymbol{x}_1 = \begin{bmatrix} -5 \\ 2 \end{bmatrix} \text{ for } \gamma_1 \text{ and}$$

$$\boldsymbol{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ for } \gamma_2 \text{ as before.}$$

Then

$$S^{-1}AS = J_A$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix},$$

where S is given by

$$S = \begin{bmatrix} -5 & -1 \\ 2 & 1 \end{bmatrix}.$$

Definition 1.3.3:

An n_i x n_i *Jordan block matrix* $J_{n_i}(\gamma_i)$ associated with an eigenvalue γ_i over a field **K** is a square matrix whose elements are 0 everywhere except on the main diagonal, where all entries equal γ_i, and in the superdiagonal where all equal 1:

$$J_{n_i}(\gamma_i) = \begin{bmatrix} \gamma_i & 1 & & \\ & \gamma_i & \ddots & \\ & & \ddots & 1 \\ & & & \gamma_i \end{bmatrix}$$

[Orte, p.120].

It is easy to check that $J_{n_i}(\gamma_i)$ has just one eigenvalue, namely γ_i with algebraic multiplicity n_i. Up to rescaling, the n_i-dimensional unit vector $\boldsymbol{e}_1$ is the only eigenvector.

Theorem 1.3.4:

Suppose the underlying field **K** is algebraically closed. Let the n x n matrix A over **K** have distinct eigenvalues $\gamma_1, \gamma_2, ---, \gamma_r$. Then there exists an invertible n x n matrix S such that

$$J_A = S^{-1}AS = \text{block diag}(J_{n_1}(\varphi_1), J_{n_2}(\varphi_2), ---, J_{n_T}(\varphi_T))$$

$$= J_{n_1}(\varphi_1) \oplus J_{n_2}(\varphi_2) \oplus --- \oplus J_{n_T}(\varphi_T), \qquad (2)$$

where the $J_{n_i}(\varphi_i)$ are n_i x n_i Jordan block matrices,

$$\sum_{i=1}^{T} n_i = \text{n},$$

and $\{\gamma_1, \gamma_2, - - - , \gamma_r\}$ = $\{\varphi_1, \varphi_2, - - - , \varphi_T\}$.

The total number T of the Jordan blocks in A is equal to the total number of the linearly independent eigenvectors of the matrix A [Laub, p. 82, Th. 9.22].

Of course, 1≤ T ≤ n and 1 ≤ n_i ≤ n in the above theorem. If T = n and n_i= 1, for i = 1, 2, - - - , n, then the Jordan canonical form (2) becomes the diagonal matrix of Proposition 1.3.1. If T = 1 and n_1 = n, then J_A itself is a Jordan block of dimension n. All possibilities between can occur.

Definition 1.3.5:

The matrix J_A in the above theorem is called the *Jordan canonical form* of the matrix A.

Definition 1.3.6:

Let A be an n x n matrix with the Jordan canonical form indicated in (2). The *elementary divisors* of the matrix A are the characteristic polynomials of the Jordan blocks of the matrix A. That is, the elementary divisor corresponding to the Jordan block $J_{n_i}(\varphi_i)$ is

$$\pi(x) = \det(J_{n_i}(\varphi_i) - xI_{n_i})$$

$$= (-1)^{n_i}(x - \varphi_i)^{n_i}$$

[Laub, p. 84].

In particular, if all the elementary divisors are linear, then J_A is a diagonal matrix. We will discuss the elementary divisors in detail in Section 1.5.

1.4 Jordan Blocks

The following algorithm determines the number of Jordan blocks and their dimensions for every eigenvalue of A. In section 1.3, we observed that, the Jordan canonical form of a matrix A is given by a direct sum of Jordan blocks each corresponding to a particular eigenvalue γ_i. We shall show how these Jordan blocks can be determined from the nullspaces of the matrices $(A - \gamma_i I_n)^j$.

<u>Jordan Blocks Algorithm 1.4.1:</u>

Suppose the underlying field **K** contains all eigenvalues for the n x n matrix A. Let γ be a particular eigenvalue, with algebraic multiplicity k.

First we solve

$$(A - \gamma I_n)\boldsymbol{v} = \mathbf{0},$$

and we let m_1 be the number of linearly independent solutions, that is,

$$m_1 = \dim(\ker(A - \gamma I_n)).$$

Suppose m_1 = k. We recall from Definition 1.1.14 that if the number of linearly independent eigenvectors (geometric multiplicity) is equal to its algebraic multiplicity, then we get a diagonal form corresponding to the eigenvalue γ, that is

$$J(\gamma) = \begin{bmatrix} \gamma & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \gamma \end{bmatrix}.$$

However, if $m_1 < k$, then we continue to solve the following homogeneous system

$$(A - \gamma I_n)^2\boldsymbol{v} = \mathbf{0}.$$

There will be m_2 linearly independent solutions where $m_2 > m_1$.

Again, if we get m_2 = k then we are done. Otherwise, solve

$$(A - \gamma I_n)^3 \boldsymbol{v} = \mathbf{0}.$$

We repeat this process until we reach

$$m_1 < m_2 < \text{- - -} < m_{N-1} < m_N = k.$$

The number N is the size of the largest Jordan block matrix associated to γ, and m_1 is the total number of blocks.

Let

$$p_1 = m_1, \ p_2 = m_2 - m_1, p_3 = m_3 - m_2, \text{- - -}, p_N = m_N - m_{N-1}.$$

Then p_i is the number of Jordan blocks of size at least i x i.

Finally we put

$q_1 = 2m_1 - m_2$, $q_2 = 2m_2 - m_1 - m_3$, - - - , $q_{N-1} = 2m_{N-1} - m_{N-2} - m_N$

and $q_N = m_N - m_{N-1}$.

Then q_s is the number of s x s Jordan blocks associated to eigenvalue γ.

After we have done this for all eigenvalues, we easily construct the Jordan canonical form J_A. Note that as long as we know the eigenvalues for A it is a fast and routine matter to compute the Jordan canonical form.

As we know that the Jordan canonical form is given by $S^{-1}AS$, now we look at how to find S.

For each γ, now order the associated Jordan blocks according to decreasing size, say

$$t_1 \geq t_2 \geq \text{- - -} \geq t_{m_1}.$$

This list therefore begins with q_N repeats of N, followed by q_{N-1} repeats of $N-1$, etc.

Then we find a vector $\boldsymbol{v}_{11}$, such that

$$(A - \gamma I_n)^{t_1} \boldsymbol{v}_{\mathbf{11}} = \mathbf{0}$$

but

$$(A - \gamma I_n)^{t_1 - 1}\boldsymbol{v}_{11} \neq \mathbf{0}.$$

Define

$$\boldsymbol{v}_{12} = (A - \gamma I_n)\boldsymbol{v}_{\mathbf{11}},\ \ \boldsymbol{v}_{13} = (A - \gamma I_n)\boldsymbol{v}_{\mathbf{12}},$$

and so on until we get $\boldsymbol{v}_{1t_1}$.

If we have one block, we are done, otherwise we can find a vector $\boldsymbol{v}_{21}$ such that

$$(A - \gamma I_n)^{t_2}\, \boldsymbol{v}_{21} = \mathbf{0},\ \ (A - \gamma I_n)^{t_2 - 1}\boldsymbol{v}_{\mathbf{21}} \neq \mathbf{0}.$$

Define

$$\boldsymbol{v}_{22} = (A - \gamma I_n)\boldsymbol{v}_{21}$$

and so on, until we get to $\boldsymbol{v}_{2t_2}$. Then if m_1 = 2, this is the end. If not, then we keep going. Eventually we get k linearly independent vectors;

$$\boldsymbol{v}_{11}, \boldsymbol{v}_{\mathbf{12}}, \text{- - -}, \boldsymbol{v}_{21}, \boldsymbol{v}_{22}, \text{- - -}, \boldsymbol{v}_{m_1 t_{m_1}}.$$

We let

$$S\gamma = (\boldsymbol{v}_{m_1 t_{m_1}}, \text{- - -}, \boldsymbol{v}_{11})$$

be the n x k matrix whose columns are these vectors in reverse order. Once we have done this for all r distinct eigenvalues, we concatenate the matrices Sγ horizontally to get an n x n matrix S. Now S will be non-singular and

$$S^{-1}AS = J_A.$$

Proof

We refer to [Laub, p. 85-89]. The key idea is that the n x n Jordan block $J_n\,(\gamma)$ has itself characteristic equation $(\gamma - x)^n = 0$, with unique eigenvalue γ having multiplicity k = n. It is easy to see that

$$(J_n\,(\gamma) - \gamma I_n)^j = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} \qquad (1)$$

has kernel spanned by $\boldsymbol{e_1}$, - - - , $\boldsymbol{e}_j$, for $j \leq n$, where $(J_n\,(\gamma) - \gamma I_n)^j = 0$ for $j \geq n$. Now suppose A has b_j blocks of size j x j for $1 \leq j \leq$ k. Then it is easy to check that

$$m_j = \dim(\ker(A - \gamma I_n)^j)$$

$$= 1b_1 + - - - + (j-1)b_{j-1} + j(b_j + - - - + b_k).$$

The assertions in the algorithm follow easily from this.■

Example 1.4.2:

Let

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & -2 & 3 \end{bmatrix}.$$

We will find the Jordan canonical form J_A of the above matrix, and find the matrix S such that

$$S^{-1}AS = J_A.$$

Solution

The characteristic equation for A is

$$(\gamma - 2)^2(\gamma - 1) = 0.$$

Its eigenvalues are given by

$$\gamma_1 = 1,\ \ \gamma_2 = \gamma_3 = 2,$$

with multiplicity of one and two respectively.

For $\gamma_2 = 2$, we first solve $(A - 2I)\boldsymbol{v} = \mathbf{0}$, by reducing

$$(A - 2I) = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 0 \\ 2 & -2 & 1 \end{bmatrix}$$

to

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

Hence, $(A - \gamma_2 I)\boldsymbol{v} = \mathbf{0}$ has one linearly independent solution, so that

$$m_1 = 1,\ \ m_1 < k = 2.$$

Next we must solve $(A - \gamma_2 I)^2\boldsymbol{v} = \mathbf{0}$, by reducing

$$\begin{bmatrix} 0 & 0 & 0 \\ -2 & 2 & -1 \\ -2 & 2 & -1 \end{bmatrix}$$

to

$$\begin{bmatrix} 1 & -1 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus $(A - \gamma_2 I)^2\boldsymbol{v} = \mathbf{0}$ has 2 linearly independent solutions and,

$$m_2 = 2 = k.$$

Now we calculate

$$p_1 = m_1 = 1$$

$$p_2 = m_2 - m_1 = 1,$$

then

$$q_1 = 2m_1 - m_2 = 0$$

$$q_2 = m_2 - m_1 = 1.$$

Hence, associated with $\gamma_2 = 2$ there is only one Jordan block of size 2 x 2 in the Jordan canonical form of A.

For γ_1, there is one Jordan block of size 1 x 1.

Hence the Jordan canonical form of the matrix A is given by

$$J_A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Now we compute the matrix S.

For $\gamma_2 = 2$, the only Jordan block that we have has size 2. We determine a vector $\boldsymbol{v}_{11}$, such that

$$(A - \gamma_2 I)^2 \boldsymbol{v}_{\mathbf{11}} = \mathbf{0} \neq (A - \gamma_2 I)\boldsymbol{v}_{11}.$$

We can take

$$\boldsymbol{v}_{11} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix},$$

then next compute

$$\boldsymbol{v}_{12} = (A - \gamma_2 I)^1 \boldsymbol{v}_{11} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

We can now let S_2 be the 3 x 2 matrix with columns $\boldsymbol{v}_{12}$ and $\boldsymbol{v}_{\mathbf{11}}$. That is

$$S_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & -2 \end{bmatrix}.$$

For

$$\gamma_1 = 1,$$

we find the vector $\boldsymbol{v}_{21}$ such that

$$(A - \gamma_1 I)^1 \boldsymbol{v}_{21} = \mathbf{0}.$$

That vector is given by

$$\boldsymbol{v}_{21} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

so that

$$S_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Concatenate S_1 and S_2 to get

$$S = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & -2 \end{bmatrix}.$$

Hence we obtain

$$S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

It has been stated before in Theorem 1.3.4 that the number of linearly independent eigenvectors is equal to the number of Jordan blocks of the matrix A. In the above example we have three vectors and two Jordan blocks for the matrix A. This motivates another definition.

Definition 1.4.3:

Let A be an n x n matrix. Then $\boldsymbol{v}$ is called the *principal vector* of degree c (sometimes called the *generalized eigenvector*) associated with an eigenvalue γ of a matrix A if and only if

$$(A-\gamma I)^{c}\boldsymbol{v} = \mathbf{0} \neq (A-\ \gamma I)^{c-1}\boldsymbol{v}.$$

[Laub, p. 85].

Below we shall describe another way of determining Jordan blocks of matrix A. We start by giving the following definitions.

Deefiniton 1.4.4:

Let $\gamma_1, \gamma_2,$ - - - , γ_r be the distinct eigenvalues of the n x n matrix A. From Theorem 1.3.4, we know that there is non-singular matrix S such that

$$
\begin{aligned}
S^{-1}AS &= J_A \\
&= \text{block diag}(J_{n_1}(\varphi_1), J_{n_2}(\varphi_2), \text{- - -}, J_{n_T}(\varphi_T)) \\
&= J_{n_1}(\varphi_1) \oplus J_{n_2}(\varphi_2) \oplus \ldots \oplus J_{n_T}(\varphi_T),
\end{aligned}
$$

and if k_i is the algebraic multiplicity of γ_i, then

$$k_1 + k_2 + \text{- - -} + k_r = \text{n}.$$

For each i let $J(\gamma_i)$ be the block matrix composed of all Jordan blocks for the eigenvalue γ_i. $J(\gamma_i)$ is called the *Jordan segment* corresponding to the eigenvalue γ_i. Now from Definition 1.1.8, the geometric multiplicity corresponding to the eigenvalue

γ_i (which in turn gives the number of the Jordan blocks in Jordan segment $J(\gamma_i)$) is given by

$$m_{i1} = \dim(\ker(\mathrm{A} - \gamma_i I_{\mathrm{n}})).$$

Let the order of the largest Jordan block in the Jordan segment $J(\gamma_i)$ be N_i, where N_i is the *index* of the eigenvalue γ_i, that is, the smallest positive integer N_i, such that

$$\mathrm{rank}(\mathrm{A} - \gamma_i I_{\mathrm{n}})^{N_i} = \mathrm{rank}(\mathrm{A} - \gamma_i I_{\mathrm{n}})^{N_i+1}.$$

Then if we let

$$m_{ij} = \dim(\ker(\mathrm{A} - \gamma_i I_{\mathrm{n}})^{j})$$

for any positive integer j and $1 \leq i \leq r$,

we get

$$m_{i1} \leq m_{i2} \leq \text{- - -} \leq m_{iN_i} = k_i.$$

From our earlier discussion we can compute from these m_{ij} the number and sizes of all Jordan blocks for γ_i. This is done in the next section.

Here we comment on a dual approach. We can use the proof of the Jordan block Algorithm 1.4.1, to find, in a slightly different way, the total number of Jordan blocks of size j x j corresponding to the eigenvalue γ_i. First consider the Jordan block $J_{n_i}(\gamma_i)$ with eigenvalue γ_i of algebraic multiplicity k_i, then

$$\mathrm{rank}(J_{n_i}(\gamma_i)^{j}) - \mathrm{rank}(J_{n_i}(\gamma_i)^{j+1}) = 0 \quad \text{for any } j \geq 1.$$

According to (1) in the proof of Jordan Block Algorithm

$$\mathrm{rank}((J_{n_i}(\gamma_i) - \gamma_i I_{n_i})^{j}) - \mathrm{rank}((J_{n_i}(\gamma_i) - \gamma_i I_{n_i})^{j+1}) = 0 \quad \text{for j} \geq k_i.$$

But

$$\text{rank}((J_{n_i}(\gamma_i) - \gamma_i I_{n_i})^j) - \text{rank}((J_{n_i}(\gamma_i) - \gamma_i I_{n_i})^{j+1}) = 1 \quad \text{for } 1 \leq \text{j} < k_i.$$

Let N_i be the size of the largest Jordan block in the Jordan segment $J(\gamma_i)$, then from the preceding above, it is easy to notice that the difference

$$d_{ij}(\gamma_i) \equiv r_{i,j\text{-}1}(\gamma_i) - r_{ij}(\gamma_i),$$

where

$$r_{ij}(\gamma_i) = \text{rank}(J(\gamma_i) - \gamma_i I)^j \quad \text{for } 1 \leq j < N_i$$

is equal to the total number of Jordan blocks of all sizes $j \leq N_i$ in $J(\gamma_i)$.

Thus the total number of Jordan blocks in Jordan segment $J(\gamma_i)$ of size j x j is equal to:

$$d_{ij}(\gamma_i) - d_{i,j+1}(\gamma_i) \equiv r_{i,j+1}(\gamma_i) - 2r_{ij}(\gamma_i) + r_{i,j\text{-}1}(\gamma_i)$$

[Meye, p. 587-590].

1.5 Elementary Divisors and Invariant Factors

This section describes the method used for determining the elementary divisors and invariant factors of a matrix A. Elementary divisors were first introduced by Weierstrass in 1868. We also relate these ideas to the Jordan canonical form.

<u>Definition 1.5.1:</u>

An n x n matrix A is *non-derogatory* if its minimal polynomial has degree n, or equivalently, if its Jordan canonical form has only one Jordan block associated with each distinct eigenvalue [Laub, p. 105].

<u>Definition 1.5.2:</u>

Suppose A is a non-derogatory n x n matrix and suppose its characteristic polynomial is given by

$$\mathrm{p}(x) = a_0 + a_1x + \text{- - -} + a_{n-1}x^{n-1} + x^n.$$

Then the matrix

$$\mathrm{A_C} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{bmatrix}$$

is called the *companion matrix* of A [Orte, p. 39].

Definition 1.5.3:

Suppose the n x n matrix A has distinct eigenvalues γ_1, γ_2, - - - , γ_r. For each eigenvalue γ_i, let

$$m_{ij} = \dim[\ker(A - \gamma_i I_n)^j], \quad j = 1, 2, ---.$$

From our earlier observations, there is an integer $N_i \geq 1$ such that

$$m_{i1} < m_{i2} < --- < m_{iN_i} = k_i,$$

where k_i is the algebraic multiplicity of γ_i. We recall that N_i is the size of the largest Jordan block in the segment $J(\gamma_i)$, and m_{i1} is the total number of blocks for γ_i. The exact number of blocks of size j is

$$q_{ij} = 2m_{ij} - m_{i,j-1} - m_{i,j+1}, \quad 1 \leq i \leq r \text{ and } 1 \leq j \leq N_i.$$

For each i we may list the block sizes in non-increasing order as

$$[t_{ij}] := [N_i, ---, N_i, N_{i-1}, ---, N_{i-1}, ---, 1, --- 1, 0, ---, 0].$$

Thus, we have in this list q_{iN_i} repeats of N_i, followed by q_{i,N_i-1} repeats ofN_{i-1}, down to q_{i1} repeats of 1. It is convenient to adjoin

$$q_{i0} := n - m_{i1},$$

repeats of 0, giving a list of length n.

Definition 1.5.4:

We may define the *invariant factors* $E_j(x)$ for A as follows.

Let

$$E_n(x) = (x - \gamma_1)^{t_{11}}(x - \gamma_2)^{t_{21}} \dots (x - \gamma_r)^{t_{r1}}.$$

One can verify that $E_n(x)$ is the minimal polynomial of the matrix A. (Recall that t_{i1} is the size of the largest Jordan block corresponding to the eigenvalue γ_i in the Jordan canonical form of the matrix A.) Now we delete from the Jordan canonical form of A one Jordan block corresponding to each factor $(x - \gamma_i)^{t_{i1}}$ of $E_n(x)$, and effectively let

$$E_{n-1}(x) = (x - \gamma_1)^{t_{12}} \dots (x - \gamma_r)^{t_{r2}}$$

be the minimal polynomial of the remaining Jordan block matrix. Again, if possible, delete one block corresponding to each factor $(x - \gamma_i)^{t_{i2}}$ and let $E_{n-2}(x)$ be the minimal polynomial of what remains and so on. In other words,

$$E_j(x) = (x - \gamma_1)^{t_{1,n+1-j}} (x - \gamma_2)^{t_{2,n+1-j}} \dots (x - \gamma_r)^{t_{r,n+1-j}}$$

(Recall that several of the exponents could well vanish.) The polynomials $E_j(x)$, j= 1, 2, - - - , n are called the *invariant factors* of A. (They are just the minimal polynomials of the series of successively deflated matrices in which certain Jordan blocks are removed at each step.)

For these invariant factors $E_j(x)$, we see that $E_j(x)$ divides $E_{j+1}(x)$. It is easy to prove that the product of the invariant factors gives the characteristic polynomial of the matrix A [HoJo, p. 154].

<u>Definition 1.5.5:</u>

Let A be a matrix of order n. Then A is similar to the matrix

$$B_{Rat} = \begin{bmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_n \end{bmatrix},$$

which is the direct sum of the companion matrices B_j of the certain uniquely determined polynomials $E_j(x)$, j = 1, 2, - - - , n, such that $E_j(x)$ divides $E_{j+1}(x)$ [BhJN, p. 410].

In fact, these polynomials $E_j(x)$ are just the invariant factors from before. If $E_j(x)$ has degree 0, as frequently happens, the block B_j is empty (0 x 0). According to Theorem 1.5.8 below, the $E_j(x)$'s can be computed in alternative fashion by rational operation over the field generated by entries of A. In other words, they can be computed without knowledge of the eigenvalues of A.

<u>Definition 1.5.6:</u>

The matrix B_{Rat} is called the *Rational canonical form* of A [Laub, p. 106].

We can give an alternate explicit description of the invariant factors after a preliminary Definition 1.5.7.

<u>Definition 1.5.7:</u>

Suppose A is an n x n matrix. We define the *determinantal divisors* $d_j(x)$ for A as follows. For j = 0, $d_0(x) \equiv 0$. For j = 1, 2, - - - , n, then $d_j(x)$ is the greatest common divisor of the determinants of the j x j submatrices of the characteristic matrix (A - xI_n) [Boch, p. 269].

Theorem 1.5.8:

Suppose A is an n x n matrix and $d_j(x)$ is the greatest common divisor of the determinants of the j x j submatrices of (A $- xI_n$), for 1$\leq j \leq$ n. Then the invariant factors $E_1(x)$, $E_2(x)$, - - - , $E_n(x)$ of the matrix (A - xI_n) are then determined by

$$E_j(x) = \frac{d_j(x)}{d_{j-1}(x)}, \quad j = 1, 2, ---, n$$

[Brua, p. 33].

Proposition 1.5.9:

Let A be a matrix of order n over an algebraically closed field and with invariant factors $E_1(x)$, $E_2(x)$, - - - , $E_n(x)$.

Recall that characteristic polynomial of the matrix A is given by

$$\text{char(A)} = \det(A - xI_n) = (x-\gamma_1)^{k_1}(x-\gamma_2)^{k_2}...(x-\gamma_r)^{k_r}$$

where γ_i are the distinct eigenvalues of the matrix A and each γ_i has algebraic multiplicity k_i. Then from the definition of the invariant factors we deduce that

$$\text{char(A)} = \det(A - xI_n) = \text{d}E_1(x)E_2(x)\ ...\ E_n(x),$$

where d is a scalar not equal to zero and $E_j(x)$ are the invariant factors of the matrix A for j = 1, 2, - - - , n. Since the invariant factors are monic, it follows that d = $(-1)^n$ and

$$(x-\gamma_1)^{k_1}(x-\gamma_2)^{k_2}...(x-\gamma_r)^{k_r} = (-1)^n E_1(x)E_2(x)\ ...\ E_n(x).$$

Moreover, since $E_j(x)$ is the divisor of $E_{j+1}(x)$, it follows from Definition 1.5.4 that;

$$E_1(x) = (x-\gamma_1)^{t_{1n}}(x-\gamma_2)^{t_{2n}}...(x-\gamma_r)^{t_{rn}}$$

$$E_2(x) = (x-\gamma_1)^{t_{1,n-1}}(x-\gamma_2)^{t_{2,n-1}}...(x-\gamma_r)^{t_{r,n-1}}$$

$$\vdots$$

$$E_n(x) = (x-\gamma_1)^{t_{11}}(x-\gamma_2)^{t_{21}}\ldots(x-\gamma_r)^{t_{r1}}$$

$$0 \le t_{in} \le t_{i,n-1} \le \text{- - -} \le t_{11}\,.$$

[LaTi, p. 266].

From Definition 1.3.6 we can notice that, each factor $(x-\gamma_i)^{t_{ij}}$ for $1 \le i \le r$ and $1 \le j \le$ n appearing in the factorization of the invariant factors in Definition 1.5.9 with $t_{ij} \ge 1$ is the elementary divisor of the matrix A.

<u>Example 1.5.10:</u>

Let A be a matrix of order six and S be a non-singular matrix such that

$$S^{-1}AS = J_A = \begin{bmatrix} J_2(\gamma_1) & & & \\ & J_2(\gamma_1) & & \\ & & J_1(\gamma_3) & \\ & & & J_1(\gamma_2) \end{bmatrix} \qquad (1)$$

$J_{n_i}(\gamma_i)$ is the Jordan block of the order n_i corresponding to the eigenvalue γ_i of the matrix A.

Now consider the matrix $J_A - xI_6$, that is

$$(J_A - xI_6) = \begin{bmatrix} J_2(\gamma_1 - x) & & & \\ & J_2(\gamma_1 - x) & & \\ & & J_1(\gamma_3 - x) & \\ & & & J_1(\gamma_2 - x) \end{bmatrix}.$$

The determinants of the Jordan blocks of the canonical form $J_A - xI_6$ are the elementary divisors of the matrix A. Then we can group together the Jordan blocks of the highest order in each eigenvalue to give the matrix G_1, then those of the next highest order to give the matrix G_2, and so on:

$$J_A = \begin{bmatrix} J_2(\gamma_1) & & & \vdots & \\ & J_1(\gamma_2) & & \vdots & \\ & & J_1(\gamma_3) & \vdots & \\ \cdots & \cdots & \cdots & \vdots & \cdots \\ & & & \vdots & J_2(\gamma_1) \end{bmatrix} = \begin{bmatrix} G_1 & \vdots & \\ \cdots & \vdots & \cdots \\ & \vdots & G_2 \end{bmatrix}.$$

Then the direct sum of the companion matrices for G_1 and G_2 gives the rational canonical form of the matrix A.

The invariant factors of A are given by

$E_1(x) = 1$

$E_2(x) = 1$

$E_3(x) = 1$

$E_4(x) = 1$

$E_5(x) = (\gamma_1 - x)^2$

$E_6(x) = (\gamma_1 - x)^2(\gamma_2 - x)(\gamma_3 - x).$

Hence, the elementary divisors of the Jordan canonical form J_A above are

$(\gamma_1 - x)^2, (\gamma_1 - x)^2, (\gamma_3 - x), (\gamma_2 - x).$

We observe that the product of the elementary divisors or invariant factors of the matrix gives the characteristic polynomial of that matrix.

If a matrix has linear elementary divisors, then that matrix has diagonal form, and if a matrix has non-linear elementary divisors, then this means that, its Jordan canonical form contains non-trivial Jordan blocks [Wilk, p. 12].

Example 1.5.11:

The square matrix A of order six with invariant factors $(x + 3)^2$ and $(x + 9)^2(x + 3)^2$ has the elementary divisors $(x + 3)^2$, $(x + 9)^2$ and $(x + 3)^2$, so that its Jordan canonical form has three Jordan blocks of order two, two of them corresponding to the eigenvalue -3 and the other one corresponding to the eigenvalue –9.

1.6 Matrices and Digraphs

In this section, we introduce directed graphs of Jordan blocks and then we use them in Chapter 2 to find the elementary divisors of a Kronecker product of Jordan blocks. Most of the material in this section can be found in [BrRy, p. 336-340] and [Brua, p. 33-44].

Definition 1.6.1:

Let A = [a_{ij}] be an n x n matrix. The *directed graph* or *digraph* D(A) = (B, C) of the matrix A is an ordered pair of two finite sets B and C, where the set C consists of some ordered pairs (i, j), or briefly ij of the elements of the set B. The elements of the set B are the *vertices* {1, 2, - - - , n} of the digraph, and the elements of the set C are the *arcs* ij from i to j whenever $i \neq j$ and $a_{ij} \neq 0$ [Brua, p. 33].

Definition 1.6.2:

A *path* δ of length $h \geq 0$ in D(A) is a sequence i_1, i_2, - - - , i_{h+1} of (h+1) vertices such that $i_1 i_2$, - - - , $i_h i_{h+1}$ are all arcs. The idea is that the paths should be unidirectional.

An α-*path* in D(A) is a set δ of vertices which can be partitioned into sets V_1, V_2, - - - , V_α, such that each V_i is the set of the vertices of a path of D(A). The *size* of the α-path δ is the number $|\delta|$ of vertices in δ. We define p_α(D) to be the largest number of the vertices in any α-path in D(A).

Thus p_1(D) is the number of vertices in its largest path in D(A), and we set p_0(D) = 0.

One can show that there is a smallest positive integer $f \leq$ n such that

$0 = p_0(\mathrm{D}) < p_1(\mathrm{D}) < \text{- - -} < p_f(\mathrm{D}) = \text{- - -} = p_n(\mathrm{D}) = \mathrm{n}.$

The sequence

$$p(\mathrm{D}) = (p_0(\mathrm{D}), p_1(\mathrm{D}), \text{- - -}, p_n(\mathrm{D}))$$

is called the *path-number sequence* of D(A). We call the integer f the *width* of the digraph D(A) and denote it by width(D(A)). That is, the width of D(A) is the smallest positive integer f such that all vertices in D(A) can be partitioned into f paths [BrRy, p. 336-337].

Example 1.6.3:

Consider the matrix

$$A = \begin{bmatrix} 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then D(A) = (B, C), where B = { 1, 2, 3, 4, } and C= {(1, 3), (2, 4), (3, 4)}. We can draw the digraph as

$$1 \rightarrow 3 \rightarrow 4 \leftarrow 2.$$

We can easily see that $p_1(\mathrm{D}) = 3$ (given by a path with vertices 1, 3 and 4). $p_2(\mathrm{D}) = 4$ (given by a path with vertices 1, 3 and 4 and another path with vertex 2). Furthermore, $f = 2$.

We shall show how the digraph D(A) of the matrix A determines the Jordan canonical form of A, in particular the width of D(A) is equal to the number of Jordan blocks in the matrix A.

Definition 1.6.4:

Let τ_{n_1} be the digraph of the Jordan block $J_{n_1}(\varphi_1)$. Its vertices are 1, 2, - - -, n_1 and its arcs are 12, - - -, $(n_1 - 1)n_1$. (Note that this structure is independent of the value of φ_1; in particular, it does not matter if $\varphi_1 = 0$.) Then for two digraphs τ_{n_1} of $J_{n_1}(\varphi_1)$ and τ_{n_2} of the block $J_{n_2}(\varphi_2)$ with the vertex sets $V_{n_1}=\{1, ---, n_1\}$ and $V_{n_2}=\{1, ---, n_2\}$, respectively, the *Cartesian product* $\tau_{n_1} \times \tau_{n_2}$ of the digraphs $\tau_{n_1}$and τ_{n_2} is the digraph whose vertex set is the Cartesian product $V_{n_1} \times V_{n_2}$ of the sets V_{n_1} and V_{n_2}, with an arc from (i_1, j_1) to (i_2, j_2) if and only if $i_1 = i_2$ and there is an arc $j_1 j_2$ in τ_{n_2} or $j_1 = j_2$ and there is an arc $i_1 i_2$ in τ_{n_1} [Brua, p. 34].

Definition 1.6.5:

The *Cartesian conjuction* $(\tau_{n_1} \wedge \tau_{n_2})$ has the vertex set $V_{n_1} \times V_{n_2}$ where there is an arc from (i_1, j_1) to (i_2, j_2) if and only if there is an arc $i_1 i_2$ in $\tau_{n_1}$and an arc $j_1 j_2$ in τ_{n_2} [Brua, p. 34].

Definittion 1.6.6:

The *Kronecker product* $\tau_{n_1} \otimes \tau_{n_2}$ of digraphs $\tau_{n_1}$and τ_{n_2} has the vertex set $V_{n_1} \times V_{n_2}$. Its arc set is the union of the arc set $\tau_{n_1} \times \tau_{n_2}$ and the arc set $\tau_{n_1} \wedge \tau_{n_2}$, that is, the arc set $(\tau_{n_1} \times \tau_{n_2}) \cup (\tau_{n_1} \wedge \tau_{n_2})$ [Brua, p. 34].

Lemma 1.6.7:

If $\varphi_1, \varphi_2 \neq 0$, then

$$\mathrm{D}(J_{n_1}(\varphi_1) \otimes J_{n_2}(\varphi_2)) = \mathrm{D}(J_{n_1}(\varphi_1)) \otimes D(J_{n_2}(\varphi_2))$$

[Brua, p. 42].

Lemma 1.6.8:

Let n_1 and n_2 be positive integers. Then for $\alpha = 1, 2, ---, \min(n_1, n_2) = \beta$,

$$p_\alpha(\tau_{n_1} \otimes \tau_{n_2}) = p_\alpha(\tau_{n_1} \times \tau_{n_2}) = \sum_{j=1}^{\alpha} [n_1 + n_2 - (2j - 1)]. \quad (1)$$

Proof

If

$$p_\alpha(\tau_{n_1} \otimes \tau_{n_2}) = p_\alpha(\tau_{n_1} \times \tau_{n_2}),$$

then an α-path δ of size $p_\alpha(\tau_{n_1} \times \tau_{n_2})$ can be partitioned into α paths such that the longest path δ_1 joins the vertex (1, 1) in the lower left corner to the vertex (n_1, n_2) in the upper right corner of digraph $\tau_{n_1} \times \tau_{n_2}$. Then the remaining paths $\delta_2, \delta_3, ---, \delta_\alpha$ can be drawn entirely above or entirely below δ_1 such that they do not share vertices. This gives us the canonical α-path of $\tau_{n_1} \times \tau_{n_2}$ which shows that

$$p_\alpha(\tau_{n_1} \otimes \tau_{n_2}) = p_\alpha(\tau_{n_1} \times \tau_{n_2}) = \sum_{j=1}^{\alpha} [n_1 + n_2 - (2j - 1)].$$

The entire proof of this Lemma can be found in [Brua, p. 34-38].■

Definition 1.6.9:

Let A be a matrix of order n and let C be $(n - \alpha)$ x $(n - \alpha)$ submatrix of A obtained by deleting α rows and α columns from A. Recall Definition 1.5.7. The greatest common divisor of the determinants of the submatrices of order $n - \alpha$ of the matrix (A - xI_n) is called the $(n - \alpha)^{\text{th}}$ determinantal divisor and is denoted by $d_{n-\alpha}(x)$, for

α = 0, 1, 2, - - -, n. We define $d_0(x)$ = 1 and there exists a positive integer y $\leq n$ and integers

$$0 = s_0 < s_1 < - - - < s_{\text{y}} = s_{\text{y+1}} = - - - = s_n = n$$

such that the degree of $d_{n-\alpha}(x)$ is $n - s_\alpha$ for α = 0, 1, - - - , n. We call the sequence

$$d(\text{A}) = (s_0, s_1, - - - , s_n)$$

the *divisor sequence* of the matrix A [BrRy, p. 337].

In chapter 2 we prove that $s_\alpha = p_\alpha$(D(A)).

Definition 1.6.10:

The *term rank* of the matrix A is the total number of nonzero entries of A having the property that no two come from the same row or column [Brua, p. 41].

Lemma 1.6.11:

Let A = $[a_{i_t j_t}]$ be a matrix of order n with zero entries in the main diagonal. Let α be a positive integer with $\alpha \leq$ width(D(A)) and let X be an α-path of the digraph D(A) of size u. Then there is a submatrix of the matrix A of order (u - α) with term rank equal to (u - α).

Proof

Since $\alpha \leq$ width(D(A)) the set X can be partitioned into exactly α paths ϑ_t joining a vertex i_t to a vertex j_t, for t = 1, 2, - - - , α. Let B be the principal submatrix of order u of the matrix A determined by the rows and the columns whose indices lie in the set X. Let C be the submatrix of B obtained by deleting columns $j_1, j_2,$ - - - , j_α and rows $i_i, i_2,$ - - - , i_α. Then C is submatrix of order (u - α) of A and C has term rank equal to (u - α) [BrRy, p. 338]■

Lemma 1.6.12:

Let T be a strictly upper triangular matrix of order n. Let σ and ε be the subsets of {1, 2, - - - , n} of size u, and suppose that the submatrix T[σ, ε] determined by the rows with indices in σ and the columns with indices in ε has term rank u. Then the complementary submatrix T[σ^*, ε^*] of order (n - u) is a strictly upper triangular matrix.

Proof

Let T[σ, ε] = [i_1, i_2 - - -, i_u|j_1, j_2, - - - , j_u] be u x u submatrix of T formed by rows i_i, i_2, - - , i_u and columns j_1, j_2, - - - , j_u. Suppose that for some w with $1 \leq w \leq u$, we have $j_w \leq i_w$. Let

$$T[\sigma, \varepsilon] = [a_{i_t j_t}] \text{ for } t = 1, ---, u.$$

Then $a_{i_t j_t}$ = 0 for i_w, - - -, i_u and j_1, - - -, j_w. Hence T[σ, ε] has an ($u - w$ + 1) x w zero submatrix and the term rank of T[σ, ε] cannot be u. Thus $j_w > i_w$ for each w = 1, 2, - - - ,u. But now it follows that the complementary submatrix T[σ^*, ε^*] of order (n - u) is a strictly upper triangular matrix [Brua, p. 41].■

<u>Lemma 1.6.13:</u>

Let T, σ and ε satisfy the assumptions of the above Lemma. Let X be a set of u non-zero elements of T[σ, ε] with no two from the same row or column. Then the arcs of the digraph D(T) corresponding to the elements of the set X can be partitioned into z ≤ (n - u) pairwise vertex disjoint paths each of which joins a vertex in ε^* to a vertex in σ^* where T[σ^*, ε^*] is a complementary submatrix of T[σ, ε]. The vertices on these paths form an (n - u)-path of the digraph D(T) of size (z + u).

The proof of this Lemma can be found in [BrRy, p. 339].

We use these combinatorial ideas in the next chapter to derive the Jordan canonical form of Kronecker product of two Jordan blocks.

Chapter 2: Jordan Canonical Form of a Kronecker Product

This chapter discusses the Jordan canonical form of a Kronecker product matrix of the two matrices A and B. The Jordan canonical form of a Kronecker product was described and the first proof of its existence ad uniqueness was given by Roth in 1934 [Roth]. The discussion on the first section represents results from Ortega [Orte] and the last section contains some results presented by Horn and Johnson [HoJo]. We also use the work of Brualdi [Brua] to derive the elementary divisors of Kronecker product of two Jordan blocks using a combinatorial derivation. We first look at the matrices with the distinct eigenvalues (that is, geometric multiplicity of each eigenvalue is one), then we discuss the more complicated Jordan canonical form of matrices with repeated eigenvalues. A large portion of this chapter deals with various aspects of eigenvalues and eigenvectors of a matrix that give rise to different Jordan structures. To determine the Jordan canonical form of a Kronecker product matrix $A \otimes B$ we need to determine the Jordan block structure of the Kronecker product matrix given by two Jordan blocks, one from matrix A and one from matrix B. After stating the results, I will illustrate the discussion by giving some examples. As an application of these ideas, at the end of this chapter we use the matrix exponential function to solve a linear system of differential equations.

2.1 Direct Sum of Jordan Blocks

In this section, we look at the contribution to the Jordan canonical form of a Kronecker product A$\otimes$B from each pair of the Jordan blocks, one from a matrix A and one from a matrix B.

Proposition 2.1.1:

If the Jordan blocks in the Jordan canonical form of a matrix A are

$J_{n_1}(\varphi_1), J_{n_2}(\varphi_2)$, - - -, $J_{n_T}(\varphi_T)$ and those of a matrix B are

$J_{m_1}(\mu_1), J_{m_2}(\mu_2)$, - - -, $J_{m_Y}(\mu_Y)$, then a Jordan canonical form of a Kronecker product matrix A$\otimes$B is the direct sum of the Jordan canonical forms of the matrices

$J_{n_i}(\varphi_i) \otimes J_{m_j}(\mu_j)$ for i = 1, 2, - - - , T and j = 1, 2, - - - , Y.

The proof of this follows from Proposition 1.2.3 part (11) [HoJo, p. 261-262].

Example 2.1.2:

Show that the Jordan canonical form of the Kronecker product matrix A$\otimes$B is given by the direct sum of the Jordan canonical forms of the Kronecker product of the Jordan blocks of the matrices A and B, if

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & -2 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}.$$

Solution

The Jordan canonical form of the matrix A is given by

$$J_A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

this Jordan canonical form has two blocks which are

$$J_1(1) = [1] \text{ and } J_2(2) = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

The Jordan canonical form of the matrix B is given by

$$J_B = J_2(2) = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix},$$

so this Jordan canonical form has one Jordan block which is itself.

Then the first Kronecker product of our Jordan blocks is given by

$$J_1(1) \otimes J_2(2) = [1] \otimes \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

The Jordan canonical form of the matrix $J_1(1) \otimes J_B$ is;

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

The second Kronecker product of our Jordan blocks is given by

$$J_2(2) \otimes J_2(2) = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \otimes \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 2 & 2 & 1 \\ 0 & 4 & 0 & 2 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

The Jordan canonical form of the matrix $J_2(2) \otimes J_2(2)$ is given below;

$$\begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Then the Jordan canonical form of the Kronecker product matrix $A \otimes B$ is given;

$$J_{(A\otimes B)} = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

If both matrices A and B are similar to diagonal forms J_A and J_B, respectively, then the direct sum of Jordan canonical forms of Kronecker product of their Jordan blocks is equivalent to $J_A \otimes J_B$. Then we have the following proposition.

<u>Proposition 2.1.3:</u>

If the number of linearly independent eigenvectors of the matrices A and B are equal to the number of eigenvalues (that is each eigenvalue has its own linearly independent eigenvector) then the matrix A and B are diagonalizable and thus we can construct the complete Jordan structure of the Kronecker product matrix $A \otimes B$ as $J_A \otimes J_B$.

<u>Proof</u>

In particular if there exists non-singular matrices S and Q such that

$$S^{-1}AS = J_A, \quad Q^{-1}BQ = J_B,$$

where J_A and J_B are diagonal forms of the matrices A and B, respectively, then we can get the diagonal form of the Kronecker product matrix $A \otimes B$ by

$$\begin{aligned}(S \otimes Q)^{-1}(A \otimes B)(S \otimes Q) &= (S^{-1} \otimes Q^{-1})(A \otimes B)(S \otimes Q) \\ &= (S^{-1}A \otimes Q^{-1}B)(S \otimes Q) \\ &= (S^{-1}AS) \otimes (Q^{-1}BQ) \\ &= J_A \otimes J_B.\end{aligned}$$

Hence if the matrices A and B are diagonalizable, so is the Kronecker product matrix $A \otimes B$, and its diagonal form matrix is given by $J_A \otimes J_B$.

Example 2.1.4:

Find the diagonal form of the Kronecker product matrix $A \otimes B$, if the matrices A and B are given by

$$A = \begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Solution

For the matrix A, there is a non-singular matrix S, given by

$$S = \begin{bmatrix} 1 & -3 \\ 1 & 2 \end{bmatrix}.$$

The matrix S is an eigenvector matrix composed of linearly independent eigenvectors of the matrix A such that

$$S^{-1}AS = J_A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}.$$

Also for the matrix B, there is non-singular matrix Q, obtained by putting the linearly independent eigenvectors of the matrix B in the columns, and that matrix Q is given by

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

Then $Q^{-1}BQ$ is the diagonal form matrix of the matrix B, given by

$$Q^{-1}BQ = J_B = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

The Kronecker product matrix of the matrices A and B is given by

$$A \otimes B = \begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix} \otimes \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & -1 & 6 & 3 \\ -1 & -2 & 3 & 6 \\ 4 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix}.$$

The non-singular matrix of our Kronecker product matrix according to the Theorem 1.2.4 is given by

$$S \otimes Q = \begin{bmatrix} 1 & -3 \\ 1 & 2 \end{bmatrix} \otimes \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & -3 & 3 \\ 1 & 1 & -3 & -3 \\ 1 & -1 & 2 & -2 \\ 1 & 1 & 2 & 2 \end{bmatrix}$$

and

$$(S \otimes Q)^{-1} = \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{-1}{5} & \frac{1}{5} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{5} & \frac{1}{5} & \frac{3}{10} & \frac{3}{10} \\ \frac{-1}{5} & \frac{1}{5} & \frac{-3}{10} & \frac{3}{10} \\ \frac{-1}{10} & \frac{-1}{10} & \frac{1}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{-1}{10} & \frac{-1}{10} & \frac{1}{10} \end{bmatrix}.$$

Hence the diagonal form matrix for the Kronecker product matrix is

$(S \otimes Q)^{-1}(A \otimes B)(S \otimes Q) = J_A \otimes J_B$

$$= \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \otimes \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -9 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

From our solution, we should note that, entries on the main diagonal are eigenvalues of the Kronecker product matrix $A \otimes B$.

2.2 Kronecker Product of Jordan Blocks

In Chapter one, we discovered that there is a relationship between the eigensystems of matrices A and B and the eigensystems of their Kronecker product A$\otimes$B. It is this relationship that helps one to compute Jordan canonical forms of large matrices. This section discusses results presented by Horn and Johnson [HoJo] and Brualdi [Brua].

<u>Theorem 2.2.1:</u>

Let $J_{n_1}(\varphi_1)$ be an n_1 x n_1 Jordan block and let $J_{n_2}(\varphi_2)$ be an n_2 x n_2 Jordan block. Then the Jordan canonical form of the Kronecker product $J_{n_1}(\varphi_1) \otimes J_{n_2}(\varphi_2)$ is given as follows:

1. If both φ_1 and φ_2 are nonzero eigenvalues, then associated with the eigenvalue $\varphi_1\varphi_2$ of Kronecker product, we have one Jordan block of size $n_1 + n_2 - (2\alpha - 1)$ for each α = 1, 2, - - - , min(n_1, n_2) = β.

   That is,

$$J_{n_1}(\varphi_1) \otimes J_{n_2}(\varphi_2) \sim \bigoplus_{\alpha=1}^{\beta} J_{n_1+n_2-2\alpha+1}(\varphi_1\varphi_2).$$

2. If $\varphi_1 \neq 0$ and $\varphi_2 = 0$, then associated with the eigenvalue $\varphi_1\varphi_2$= 0 of $J_{n_1}(\varphi_1) \otimes J_{n_2}(0)$, there are n_1 Jordan blocks of size n_2.

   That is,

$$J_{n_1}(\varphi_1) \otimes J_{n_2}(0) \sim \bigoplus_{i=1}^{n_1} J_{n_2}(0).$$

3. If $\varphi_1 = 0$ and $\varphi_2 \neq 0$, then associated with the eigenvalue $\varphi_1\varphi_2 = 0$ there are n_2 Jordan blocks of size n_1.

   That is,

$$J_{n_1}(0) \otimes J_{n_2}(\varphi_2) \sim \bigoplus_{i=1}^{n_2} J_{n_1}(0).$$

4. If $\varphi_1 = \varphi_2 = 0$, then associated with the eigenvalue zero of the Kronecker product there are two Jordan blocks of each size

   $\alpha = 1, 2, ---, \min(n_1, n_2) - 1$ and also there are $\rho = (n_1 + n_2 - 2\beta + 1)$ Jordan blocks of size $\min(n_1, n_2) = \beta$.

   That is,

$$J_{n_1}(0) \otimes J_{n_2}(0) \sim \bigoplus_{\alpha=1}^{\beta-1} (J_\alpha(0) \oplus J_\alpha(0)) \oplus \{\bigoplus_{i=1}^{\rho} J_\beta(0)\}$$

   [HoJo, p. 262-263, Th. 4.3.17].

Example 2.2.2:

Evaluate the Jordan canonical form of the Kronecker product of the following matrices

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & -1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

Solution

The matrix A has the eigenvalues $\varphi_i = 2$, for $i = 1, 2, 3$. This repeated eigenvalue is associated with the Jordan block of size three, that is

$$J_A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

The matrix B has two eigenvalues, $\mu_1 = 0$ and $\mu_2 = 5$. Hence this matrix has the following diagonal form

$$J_B = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}.$$

According to Theorem 1.2.4, some of the eigenvalues of the Kronecker product matrix $A \otimes B$ are given by

$$\varphi_i \mu_1 = 0, \text{ for } i = 1, 2, 3.$$

Here we have $n_1 = 3$ and $m_1 = 1$, then according to assertion 2 of the Theorem 2.2.1, associated with an eigenvalue zero of the Kronecker product, there are three Jordan blocks of size one.

For $\varphi_i \mu_2 = 10$, with $i = 1, 2, 3$,

we use the first assertion, that is, associated with the above eigenvalue of the Kronecker product matrix $A \otimes B$, there is one Jordan block of size $(n_1 + m_2 - (2\alpha - 1))$, for $n_1 = 3$, $m_2 = 1$ and $\alpha = 1, 2, ---, \min(n_1, m_2)$.

Therefore, we have one possibility, that is for $\alpha = 1$, which gives the Jordan block of size

$$(n_1 + m_2) - (2\alpha - 1) = 3.$$

Hence, an eigenvalue 10 is associated with the Jordan block of size three. Now we can construct the Jordan canonical form of the Kronecker product $A \otimes B$ as follows:

$$J_{(A\otimes B)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 & 1 & 0 \\ 0 & 0 & 0 & 0 & 10 & 1 \\ 0 & 0 & 0 & 0 & 0 & 10 \end{bmatrix}.$$

Example 2.2.3:

Find the Jordan canonical form of the Kronecker product matrix $A \otimes B$ if the matrices A and B are given by

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

Solution

The Jordan canonical forms of the matrices A and B are J_A and J_B, respectively, and they are given below:

$$J_A = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \qquad J_B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The matrix A has the eigenvalues $\varphi_1 = 0$ and $\varphi_2 = 2$, and the matrix B has the eigenvalues $\mu_i = 1$, for $i = 1, 2$.

We obtain the first two eigenvalues of the Kronecker product as

$$\varphi_1\mu_i = 0, \text{ for } i = 1, 2.$$

Then according to the third assertion of Theorem 2.2.1, associated with the eigenvalue zero of the Kronecker product there are two Jordan blocks of size one, since m_1 = 2 and n_1 = 1.

The other two eigenvalues of the matrix A⊗B are given by

$$\varphi_2\mu_i = 2, \text{ for } i = 1, 2.$$

Then we use the first assertion to obtain the Jordan block associated with an eigenvalue 2. As we know we should have one Jordan block of size (n_2 + m_1 – (2α – 1)), for n_2 = 1, m_1 = 2 and α = 1.

Hence, an eigenvalue 2 is associated with the Jordan block of size 2. Now we obtain the Jordan canonical form of the Kronecker product matrix A⊗B as follows:

$$J_{(A\otimes B)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

The following Theorem asserts that, the divisor sequence d(A) of the matrix A is equal to the path-number sequence p(D) of the digraph D(A).

<u>Theorem 2.2.4:</u>

Let A be a strictly upper triangular matrix of order n, and let α be an integer with $\alpha \leq$ width(D(A)). Then the degree (n$-s_\alpha$) of the determinantal divisor $d_{n-\alpha}(x)$ of the matrix A satisfies

$$n - s_\alpha = n - p_\alpha(D(A)).$$

Proof

It follows from Lemma 1.6.11, that there is a submatrix of A of order p_α(D(A)) $-\ \alpha$ with a non-zero term in its determinant expansion. Let C be a submatrix of order n $-\ \alpha$ of xI_n $-$ A. Then by Lemma 1. 6. 12, the lowest degree of a non-zero term in the determinant expansion of C is at least n $-\ p_\alpha$(D(A)). Then

$$n - s_\alpha \leq n - p_\alpha(D(A)).$$

Now let

$$B = (xI_n - A)[\sigma, \varepsilon]$$

be a submatrix of order n $-\ \alpha$ of xI_n $-$ A such that the det(B) has a non-zero term of degree n $-\ s_\alpha$. Then, there exists a set ϑ of size n $-\ s_\alpha$ with $\vartheta \subseteq \sigma \cap \varepsilon$ such that there is a non-zero term in the determinant expansion of the submatrix A[σ- ϑ, ε- ϑ] of order $s_\alpha - \alpha$ of the strictly upper triangular complementary submatrix A[ϑ^*, ϑ^*] of order s_α. By Lemma 1.6.13 where T is strictly upper triangular matrix A[ϑ^*, ϑ^*] of order s_α and

$$u = s_\alpha - \alpha,$$

we conclude that there exists z-path X of size $s_\alpha - \alpha$ + z in the digraph D(A) where z $\leq \alpha$. There are (α $-$ z) + (n $-\ s_\alpha$) vertices not on this z-path. Adding any α $-$ z vertices to X we obtain an α-path of size s_α. Hence p_α(D(A)) $\geq s_\alpha$ and thus

$$n - p_\alpha(D(A)) \leq n - s_\alpha.$$

Then we conclude that,

$$p_\alpha(D(A)) = s_\alpha, \quad \text{for all } \alpha = 1, 2, \text{- - -}, n$$

[BrRy, p. 339-340, Th. 9.8.4].■

Let S and Q be non-singular matrices which transform the n x n matrix A and the m x m matrix B to their respective Jordan canonical forms, J_A and J_B, that is

$$S^{-1}AS = J_A \text{ and } Q^{-1}BQ = J_B \tag{i}$$

where

$$J_A = \begin{bmatrix} J_{n_1}(\varphi_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\varphi_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & J_{n_T}(\varphi_T) \end{bmatrix},$$

$$J_B = \begin{bmatrix} J_{m_1}(\mu_1) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\mu_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & J_{m_Y}(\mu_Y) \end{bmatrix} \tag{ii}$$

and we have the Jordan blocks given by

$$J_{n_i}(\varphi_i) = \begin{bmatrix} \varphi_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \varphi_i & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \varphi_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \varphi_i \end{bmatrix},$$

$$J_{m_j}(\mu_j) = \begin{bmatrix} \mu_j & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu_j & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \mu_j & 1 \\ 0 & 0 & 0 & \cdots & 0 & \mu_j \end{bmatrix}$$

for i = 1, 2, - - -, T and j = 1, 2, - - -, Y.

Consequently $A - xI_n$ has the elementary divisors $(\varphi_i - x)^{n_i}$, and $B - xI_m$ has the elementary divisors $(\mu_j - x)^{m_j}$ [Roth, p. 462-463].

From Proposition 2.1.1, the Jordan canonical form of the Kronecker product matrix A $\otimes$ B is similar to the Jordan canonical form of the following matrix;

$$(S \otimes Q)^{-1}(A \otimes B)(S \otimes Q) = \bigoplus_{i=1}^{T} \bigoplus_{j=1}^{Y} J_{n_i}(\varphi_i) \otimes J_{m_j}(\mu_j).$$

The Jordan canonical form of the Kronecker product $J_{n_i}(\varphi_i) \otimes J_{m_j}(\mu_j)$ can be easily determined using Theorem 2.2.1.

It is very important to note that, the geometric multiplicity of an eigenvalue is equal to the number of elementary divisors associated with it and the algebraic multiplicity of an eigenvalue is equal to the sum of the degrees of all the elementary divisors associated with it.

With the above definition we can restate the Theorem 2.2.1 as follows

Theorem 2.2.5:

Let A be a matrix of order n and let B be a matrix of order m, and let A have an elementary divisor $(\varphi_1 - x)^{n_1}$ corresponding to a Jordan block $J_{n_1}(\varphi_1)$ of order n_1 in the Jordan canonical form of the matrix A and let B have an elementary divisor $(\mu_1 - x)^{m_1}$ corresponding to a Jordan block $J_{m_1}(\mu_1)$ of order m_1 in the Jordan canonical form of the matrix B, and if $\beta = \min(n_1, m_1)$, then the elementary divisors of the matrix $((J_{n_1}(\varphi_1) \otimes J_{m_1}(\mu_1)) - xI_{n_1 m_1})$ in the Jordan canonical form of A $\otimes$ B are as follows:

(1) $(\varphi_1\mu_1 - x)^{n_1+m_1-2\alpha+1}$ for α = 1, 2, - - -, β, if $\varphi_1\mu_1 \neq 0$.

(2) x^{n_1} occurring m_1 times, if $\varphi_1 = 0$ and $\mu_1 \neq 0$.

(3) x^{m_1} occurring n_1 times, if $\varphi_1 \neq 0$ and $\mu_1 = 0$.

(4) x, x^2, - - -, $x^{\beta-1}$ each occurring twice and x^{β} occurring $(n_1 + m_1 - 2\beta + 1)$ times, if $\varphi_1 = \mu_1 = 0$.

<u>Proof</u>

The elementary divisors of $((A \otimes B) - xI_{nm})$ are identical with those of $((J_A \otimes J_B) - xI_{nm})$ since

$$(S \otimes Q)^{-1}\{(A \otimes B) - xI_{nm}\}(S \otimes Q) = (J_A \otimes J_B) - xI_{nm}.$$

Then by Proposition 2.1.1, the elementary divisors of $((A \otimes B) - xI_{nm})$ are identical with the aggregate of those of the TY matrices

$(J_{n_i}(\varphi_i) \otimes J_{m_j}(\mu_j)) - xI_{n_i m_j}$, for i = 1, 2, - - - , T and j = 1, 2, - - -,Y

each taken independently, where $J_{n_i}(\varphi_i)$ is the n_i x n_i Jordan block in the Jordan canonical form of the matrix A, and $J_{m_j}(\mu_j)$ is the m_j x m_j block in the Jordan canonical form of the matrix B.

If $\mu_1 \neq 0$, $\varphi_1 \neq 0$ and $J_{n_1}(\varphi_1)$ is a non-singular matrix, then

$$((J_{n_1}(\varphi_1) \otimes J_{m_1}(\mu_1)) - xI_{n_1 m_1}) \qquad \text{(iii)}$$

is the block matrix with

$$\varphi_1 J_{m_1}(\mu_1) - xI_{m_1} \tag{iv}$$

appearing n_1 times along the main diagonal and $J_{m_1}(\mu_1)$ appearing $(n_1 - 1)$ times along the super diagonal. Let

$$J_{n_1}(\varphi_1) \otimes J_{m_1}(\mu_1) = \text{H and } n_1 m_1 = l.$$

Then, according to Definition 1.6.9, the determinantal divisor $d_{n-\alpha}(x)$ is given by $(\varphi_1\mu_1 - x)^{l-s_\alpha}$ for α = 1, 2, - - - , min(n_1, m_1) = β. It follows from Theorem 2.2.4 that

$$s_\alpha = p_\alpha(\text{D(H)})$$

such that

$$d_{n-\alpha}(x) = (\varphi_1\mu_1 - x)^{l-p_\alpha(\text{D(H)})}.$$

Then, it follows from Lemma 1.6.8 that,

$$p_\alpha(\text{D(H)}) = \sum_{j=1}^{\alpha} [n_1 + m_1 - (2j-1)].$$

Hence

$$\pi(x) = \frac{d_{n-\alpha+1}(x)}{d_{n-\alpha}(x)} = (\varphi_1\mu_1 - x)^{n_1+m_1-2\alpha+1}$$

is an elementary divisor of the matrix H by Definition 1.5.8. Since the product of the elementary divisors of the matrix H gives the characteristic polynomial of H which has the degree $n_1 m_1$ and

$$\sum_{\alpha=1}^{\beta} [n_1 + m_1 - (2\alpha - 1)] = n_1 m_1 \text{ , for } \beta = \min(n_1, m_1),$$

then we can say that, these are all the elementary divisors of the matrix H. Therefore, the elementary divisors of (iii) are given by

$$\pi(x) = (\varphi_1\mu_1 - x)^{n_1+m_1-2\alpha+1} \quad \text{for } \alpha = 1, 2, \text{- - -}, \min(n_1, m_1).$$

This proves the first assertion of our Theorem.

If $\varphi_1 = 0$, then we can easily notice that the elementary divisors of the matrix $\varphi_1 J_{m_1}(\mu_1) - xI_{m_1}$ are x occurring m_1 times, but $\varphi_1 J_{m_1}(\mu_1) - xI_{m_1}$ is appearing n_1 times along the main diagonal of the matrix (iii), hence if $\varphi_1 = 0$, then the elementary divisors of the matrix (iii) are x^{n_1} occurring m_1 times . This proves the second assertion of our Theorem.

If $\mu_1 = 0$ and $\varphi_1 \neq 0$, then the Kronecker product matrix H is the nilpotent matrix such that its m_1th power is zero but no lower power vanishes. Hence, in this case, the matrix (iii) has only elementary divisors of the form x^t, where $t \leq m_1$. But since $\varphi_1 \neq 0$, then the matrix in (iii) has non-singular minor of the order $(m_1 - 1)n_1$ which is independent of x, hence by Rank Theorem 1.1.13, at most

$$n_1 m_1 - (m_1 - 1)n_1 = n_1$$

elementary divisors can be powers of x. The determinant of the matrix in (iii) is $x^{n_1 m_1}$ if $\mu_1 = 0$, consequently n_1 elementary divisors must be x^{m_1}. This establishes the third case under the theorem.

Let $\varphi_1 = \mu_1 = 0$, then if we remove the first columns and the last rows of the blocks $J_{n_1}(\varphi_1)$ and $J_{m_1}(\mu_1)$, respectively, we get $(n_1 - 1)$ ones and $(m_1 - 1)$ ones on the diagonals of the matrices $J_{n_1-1}(\varphi_1)$ and $J_{m_1-1}(\mu_1)$, respectively, hence H has $(n_1 - 1)(m_1 - 1)$ ones in the m_1th diagonal above its super diagonal and all the remaining entries are zeros. From the non-singular minor matrix of the order $(n_1 -$

$1)(m_1 - 1)$, which is independent of x, and by the Rank Theorem 1.1.13, the matrix in (iii) has

$$n_1 m_1 - (n_1 - 1)(m_1 - 1) = n_1 + m_1 - 1$$

elementary divisors which are powers of x. Now H is nilpotent of the degree $\min(n_1, m_1)$ for in the present case

$$(J_{n_1}(\varphi_1))^{\beta} = 0$$

or

$$(J_{m_1}(\mu_1))^{\beta} = 0, \quad \text{where } \beta = \min(n_1, m_1)$$

but for $g < \min(n_1, m_1)$, then

$$(J_{n_1}(\varphi_1))^{g} \neq 0 \text{ and } (J_{m_1}(\mu_1))^{g} \neq 0.$$

Consequently the matrix in (iii) has $(n_1 + m_1 - 1)$ elementary divisors x^c, whose degree c does not exceed $\beta = \min(n_1, m_1)$ and must at least one of degree β. We may, therefore, assume that the matrix in (iii) has the elementary divisors x^c, occurring v_c times, for $c = 1, 2, - - -, \beta$. Hence its determinant is $(-x)$ raised to the power,

$$v_1 + 2v_2 + - - - + \beta v_{\beta} = n_1 m_1. \tag{v}$$

Now let the Jordan canonical form of the Kronecker product matrix

$J_{n_1}(\varphi_1) \otimes J_{m_1}(\mu_1)$ be $J_{n_1 m_1}(\varphi_1 \mu_1)$.

Then $(J_{n_1 m_1}(\varphi_1 \mu_1))^c$ has

$$v_{c+1} + 2v_{c+2} + - - - + (\beta - c)v_{\beta}$$

ones in the c^{th} diagonal above its main diagonal and zeros else where, and

$(J_{n_1}(\varphi_1) \otimes J_{m_1}(\mu_1))^c$ has $(n_1 - c)(m_1 - c)$ ones in the $c(m_1 + 1)^{\text{th}}$ diagonal above its main diagonal and zeros else where. Since these two matrices are of the same rank for all values of c, we have

$v_{c+1} + 2v_{c+2} + \text{- - -} + (\beta - c)v_\beta = (n_1 - c)(m_1 - c)$, for $c = 1, 2, \text{- - -}, \beta - 1$. (vi)

Thus combining (v) and (vi) we obtain the triangular system of equations

$$v_1 + 2v_2 + \text{- - -} + \beta v_\beta = n_1 m_1$$

$$v_2 + 2v_3 + \text{- - -} + (\beta - 1)v_\beta = (m_1 - 1)(n_1 - 1)$$

$$\vdots$$

$$(\beta - (\beta - 1)v_\beta = \{(n_1 - (\beta - 1))(m_1 - (\beta - 1))\}$$

which has the solution

$$v_{n_1} = m_1 - n_1 + 1.$$

If we let

$$\min(n_1, m_1) = \beta = n_1$$

and

$$v_c = 2, \quad c = 1, 2, \text{- - -}, n_1 - 1.$$

Thus in the case when

$$\varphi_1 = \mu_1 = 0$$

the matrix in (iii) has $(m_1 + n_1 - 1)$ elementary divisors as stated by the above Theorem [Roth, p. 463-464, Th. 1].■

2.3 Matrix Exponential Function

In the following section, we use the Kronecker products in transforming matrix equations into corresponding matrix-vector equations. Then, we use the Jordan canonical form to derive solutions of differential equations.

Definition 2.3.1:

Let A_j denote the j^{th} column of an m x n matrix A. Then *mn column vector* of A is defined as

$$v(A) = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$$

[HoJo, p. 244].

Lemma 2.3.2:

Let A be an m x n matrix, B be an n x p matrix and C be p x q matrix for which the matrix ABC is defined, then

$$v(ABC) = (C^{T} \otimes A)v(B).$$

Proof

For a given matrix R, let R_b denote the b^{th} column of R. Let

$$C = [c_{ij}].$$

Then

$$(ABC)_b = A(BC)_b$$
$$= ABC_b$$
$$= A\{\sum_{i=1}^{p} c_{ib} B_i\}$$
$$= [c_{1b}A \;\; c_{2b}A \; . \; . \; . \; c_{pb}A]v(B)$$
$$= (C_b{}^{\mathrm{T}} \otimes A)v(B).$$

Thus,

$$v(ABC) = \begin{bmatrix} C_1^T \otimes A \\ \vdots \\ C_q^T \otimes A \end{bmatrix} v(B).$$

But this product is just $(C^{\mathrm{T}} \otimes A)v(B)$ since the transpose of a column of C is a row of C^{T} [HoJo, Lem 4.3.1, p. 254-255].■

Definition 2.3.3:

For any matrix X, its *matrix exponential function* is given by

$$e^{tX} = \sum_{d=0}^{\infty} \frac{(tX)^d}{d!}.$$

This converges for any matrix as shown in James M. Ortega [Orte, p. 177].

Proposition 2.3.4:

For matrices A and B of order n and if A and B commutes, that is, if $AB = BA$, then

$$e^{t(A+B)} = e^{tA}e^{tB} = e^{tB}e^{tA}.$$

Proof

From Definition 2.3.3,

$$e^{t(A+B)} = I_n + t(A+B) + \frac{t^2}{2!}(A+B)^2 + \ldots$$

and

$$e^{tA}e^{tB} = (I_n + tA + \frac{t^2}{2!}A^2 + \ldots)(I_n + tB + \frac{t^2}{2!}B^2 + \ldots)$$

while

$$e^{tB}e^{tA} = (I_n + tB + \frac{t^2}{2!}B^2 + \ldots)(I_n + tA + \frac{t^2}{2!}A^2 + \ldots).$$

Comparing coefficients of like powers of t in the first equation and the second or third equation gives us the desired results [Laub, p. 110].■

Definition 2.3.5:

Let A_1 and A_2 be matrices of order 2. Then (A_1, A_2) is a *matrix resolution of the identity* if,

$$(A_1)^2 = A_1, \quad (A_2)^2 = A_2$$

$$A_1A_2 = A_2A_1 = 0$$

$$A_1 + A_2 = I_2 \tag{1}$$

where I_2 is 2 x 2 identity matrix.

Suppose A is a diagonalizable matrix of order 2 and γ_1 and γ_2 are eigenvalues of the matrix A, then the *spectral decomposition* of A is given by

$$A = \gamma_1 A_1 + \gamma_2 A_2. \tag{2}$$

Now we can show how to obtain the matrices A_1 and A_2. From (1), we can replace A_2 by $(I_2 - A_1)$ in (2) obtaining

$$A_1 = (\gamma_1 - \gamma_2)^{-1}(A - \gamma_2 I_2).$$

With the same argument we can obtain

$$A_2 = (\gamma_2 - \gamma_1)^{-1}(A - \gamma_1 I_2).$$

The matrices A_1 and A_2 are called the *spectral bases* of A and the set of distinct eigenvalues of A is called the *spectrum*.

In general if A is a diagonalizable matrix of order n and $(\gamma_1, \gamma_2, \text{- - -}, \gamma_n)$ are its eigenvalues. Then the spectral bases of the matrix A are given by

$$Ai = \prod_{\substack{j=1 \\ j\neq i}}^{n} (\gamma_i - \gamma_j) \prod_{\substack{j=1 \\ j\neq i}}^{n} (A - \gamma_j I_n), \text{ for } i = 1, 2, \text{- - -}, n$$

[Meye, p. 520].

<u>Proposition 2.3.6:</u>

Let A be a diagonalizable matrix whose eigenvalues are $(\gamma_1, \gamma_2, \text{- - -}, \gamma_n)$, then the *exponential function* of A is given by

$$e^{At} = e^{\gamma_1 t} A_1 + e^{\gamma_2 t} A_2 + \text{- - -} + e^{\gamma_n t} A_n$$

where the Ai are the spectral bases. The proof is in [Meye, p. 530].

Definition 2.3.7:

Theorem 1.3.4 stated that, for any matrix A, there is non-singular matrix that transforms A into its Jordan canonical form, that is

$$S^{-1}AS = J_A = \text{block diag}(J_{n_1}(\varphi_1), J_{n_2}(\varphi_2), \text{- - -}, J_{n_T}(\varphi_T))$$

$$= J_{n_1}(\varphi_1) \oplus J_{n_2}(\varphi_2) \oplus \text{- - -} \oplus J_{n_T}(\varphi_T),$$

where $J_{n_i}(\varphi_i)$ are the n_i x n_i Jordan blocks associated with the eigenvalue φ_i, it then follows from Definition 2.3.3 that the *exponential function* of A is given by

$$e^{At} = Se^{tJ_A}S^{-1} = S\begin{bmatrix} \ddots & & \\ & e^{J_{n_i}(\varphi_i)t} & \\ & & \ddots \end{bmatrix}S^{-1}.$$

If we let N^j be the nilpotent matrix associated with the eigenvalue φ_i for i = 1, 2, - - - , T and for j = 0, 1, 2, - - - , $n_i - 1$, where n_i = index(φ_i).

Again by Definition 2.3.3

$$e^{J_{n_i}(\varphi_i)t} = e^{\gamma_i It + Nt}$$

$$= e^{\varphi_i t}\left(I + tN + \frac{1}{2!}t^2N^2 + \text{- - -} + \frac{1}{(n_i-1)!}t^{n_i-1}N^{n_i-1}\right)$$

$$= e^{\varphi_i t}\sum_{j=0}^{n_i-1}\frac{1}{j!}t^jN^j$$

$$= e^{\varphi_i t}\begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & \vdots & 0 & 1 & t \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

[Laub, p. 114-115].

Example 2.3.8:

Find the solution of the 2 x 2 matrix differential equation

$$\frac{dX}{dt} = AXB$$

with

$$A = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}, \quad B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \text{ and } a \neq b.$$

Solution

The eigenvalues of the matrix A are 1 and -2. Let

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix},$$

then

$$\frac{dX}{dt} = AXB$$

$$\frac{d}{dt}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

$$= \begin{bmatrix} 10ax_{11} - 18ax_{21} & 10bx_{12} - 18bx_{22} \\ 6ax_{11} - 11ax_{21} & 6bx_{12} - 11bx_{22} \end{bmatrix}.$$

Then according to Definition 2.3.1, we get

$$\frac{d}{dt}\begin{bmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 10ax_{11} - 18ax_{21} \\ 6ax_{11} - 11ax_{21} \\ 10bx_{12} - 18bx_{22} \\ 6bx_{12} - 11bx_{22} \end{bmatrix}.$$

This matrix equation can be written as

$$\frac{dv(X)}{dt} = (B^T \otimes A)v(X) \tag{3}$$

with

$$(B^T \otimes A) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \otimes \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}$$

$$= \begin{bmatrix} 10a & -18a & 0 & 0 \\ 6a & -11a & 0 & 0 \\ 0 & 0 & 10b & -18b \\ 0 & 0 & 6b & -11b \end{bmatrix}.$$

The solution to equation (3) is

$$v(X) = e^{(B^T \otimes A)t} v(X(0))$$

such that

$$X(t) = e^{At} X(0)\, e^{Bt}.$$

Then by Definition 2.3.5 the spectral decompositions are given by

$$A = A_1 - 2A_2, \quad B = aB_1 + bB_2$$

and the spectral bases are given by

$$A_1 = \frac{1}{3}(A + 2I_2) = \begin{bmatrix} 4 & -6 \\ 2 & -3 \end{bmatrix}$$

$$A_2 = \frac{-1}{3}(A - I_2) = \begin{bmatrix} -3 & 6 \\ -2 & 4 \end{bmatrix}$$

$$B_1 = (a - b)^{-1}(B - bI_2) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$B_2 = (b - a)^{-1}(B - aI_2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

I_2 is just 2 x 2 identity matrix.

Then from Definition 2.3.6 we can immediately write down the solution as

$$X(t) = (e^{(1+a)t} A_1 + e^{(-2+a)t} A_2) X(0)\, B_1 + (e^{(1+b)t} A_1 + e^{(-2+b)t} A_2) X(0) B_2. \tag{4}$$

Let

$$X(0) = \begin{bmatrix} x_{11}(0) & x_{12}(0) \\ x_{21}(0) & x_{22}(0) \end{bmatrix}.$$

Then from (4) we get

$$X(t) = (e^{(1+a)t}A_1 + e^{(-2+a)t}A_2)\begin{bmatrix} x_{11}(0) & 0 \\ x_{21}(0) & 0 \end{bmatrix} + (e^{(1+b)t}A_1 + e^{(-2+b)t}A_2)\begin{bmatrix} 0 & x_{12}(0) \\ 0 & x_{22}(0) \end{bmatrix}.$$

General Solution

Let

$$\frac{dX}{dt} = AXB$$

which can be transformed into

$$\frac{dv(X)}{dt} = \left(B^T \otimes A\right)v(X),$$

this has the general solution

$$v(X) = e^{(B^T \otimes A)t}v(X(0)). \tag{5}$$

Application of Main Theorem 2.2.5:

Let the Jordan canonical form for B^T be n x n Jordan block $J_n(\varphi)$ associated with non-zero eigenvalue φ and let the Jordan canonical form for A be m x m Jordan block $J_m(\mu)$ corresponding to non-zero eigenvalue μ.

Let

$$J_n(\varphi) \otimes J_m(\mu) = C,$$

then by Theorem 2.2.1, there exists a non-singular matrix S, such that

$$\mathrm{S}^{-1}C\mathrm{S} = J_C = \bigoplus_{\alpha=1}^{\beta} J_{n+m-2\alpha+1}(\varphi\mu) \tag{6}$$

where J_C is the Jordan canonical form for C.

Then by Definition 2.3.7, the solution of (5) is given by

$$X(t) = \mathrm{S}e^{tJ_C}\mathrm{S}^{-1}X(0).$$

Let min(n, m) = β = n, then from (6) and Definition 2.3.7,

$$\mathrm{S}e^{tJ_C}\mathrm{S}^{-1} = \mathrm{S}\begin{bmatrix} e^{J_{n+m-1}(\varphi\mu)t} & & \\ & \ddots & \\ & & e^{J_{m-n+1}(\varphi\mu)t} \end{bmatrix}\mathrm{S}^{-1}$$

$$= e^{\varphi\mu t}\,\mathrm{S}\{\bigoplus_{\alpha=1}^{n}\sum_{j=0}^{n+m-2\alpha}(\frac{1}{j!}N_{n+m-2\alpha+1}^{j}t^{j})\}\mathrm{S}^{-1},$$

where $N_{n+m-2\alpha+1}$ is a nilpotent matrix of order $n + m - 2\alpha + 1$ for each α = 1, - - -, n.

Bibliography

1. [Barn] S. Barnett, “Matrix Differential Equations and Kronecker Products”, Journal on Applied Mathematics, Vol. **24**, Number 1, January 1973, p. 1-5.

2. [Boch] M. Bocher, “Introduction to Higher Algebra”, Macmillan Company, New York (1907), p. 262-273.

3. [BhJN] P.B. Bhattachrya, S.K. Jain, S.R. Nagpaul, “Basic Abstract Algebra”, Cambridge University Press, New York(1994), p. 417-425.

4. [Bran] L. Brand, “The Companion Matrix and its Properties”, Mathematical Association of America, Vol. **71**, Number 6, July 1964, p. 629-634.

5. [BrRy] R. A. Brualdi, H. J. Ryser, “Combinatorial Matrix Theory”, Cambridge University Press, New York (1991), p. 335-340.

6. [Brua] R. A. Brualdi, “Combinatorial Verification of the Elementary Divisors of Tensor Products”, Linear Algebra and its Applications, Vol.**71**, 1985, p. 31-47.

7. [HoJo] R.A. Horn, C.R. Johnson, “Topics in Matrix Analysis”, Cambridge University Press, New York (1991), p. 242-263.

8. [Laub] A.J. Laub, “Matrix Analysis for Scientist and Engineers”, Society for Industrial and Applied Mathematics, California (2005), p. 140- 142.

9. [LaTi] P. Lancaster, M. Tismenetsky, “The Theory of Matrices”, Academic Press, New York (1985), p. 259-274.

10. [LiSt] W.S. Li, E. Strouse, “Reflexivity of Tensor Products of Linear Transformation”, Proceedings of the American Mathematical Society, Vol. **123**, Number 7, July 1995, p. 2021-2029.

11. [MaRo] M. Marcus, H. Robinson, “Elementary Divisors of Tensor Products”, Association for Computing Machinery”, Vol. **18**, Number 1, January 1975, p. 36-39.

12. [Meng] W.O. Menge, “On the Rank of the Product of Certain Square Matrices”, Bulletin of the American Mathematical Society, Vol. **38**, Number 2, 1932, p. 88-94.

13. [Meye] C.D. Meyer, “Matrix Analysis and Applied Linear Algebra”, Society for Industrial and Applied Mathematics, Philadelphia (2000), p. 573-598.

14. [Orte] J.M. Ortega, “Matrix Theory (A second course)”, Plenum Press, New York and London (1989), p. 117-129.

15. [Pool] D. Poole, “Linear Algebra: A Modern Introduction”, Thomson Brooks/Cole, Canada(2006), p. 202-203.

16. [Roth] W.E. Roth, “On Direct Product Matrices”, Bulletin of the American Mathematical Society, Vol.**40**, Number 6, 1934, p. 461-468.

17. [Stra] G. Strang, “Introduction to Linear Algebra”, Wellesley-Cambridge Press, Wellesley (1993), p. 248-249.

18. [Wilk] J.H. Wilkson, “The Algebraic Eigenvalue Problem”, Oxford University Press, London (1965), p. 10-39.

Curriculum Vitae

Candidate's full name: Pako Ramasu

Education:

2002-2006: University of Botswana, Gaborone, Botswana

B. Ed (Mathematics)

2007-2008: University of New Brunswick, Fredericton NB, Canada

Graduate in Department of Mathematics & Statistics

Thesis: The Jordan Canonical Form of a Kronecker Product

Supervisor: Dr. Colin Ingalls

MSc (Mathematics) will be awarded May 2009.