Chapter 4

diagonalization and Jordan form of the companion matrix

4.1 introduction

It may seem that, in chapter 3, we have effortlessly solved both the homogenous and the inhomogeneous scalar differential and difference equation of degree \( N \) by recasting these, using the companion matrix \( C \), as vector differential or difference equations of degree 1. There is a catch, directly computation of \( \exp(tC) \) and \( C^n \) which lie at the heart of these solutions, is not easy. However these computations are possible if the companion matrix \( C \) can be diagonalized, i.e. if \( C \) has a full complement of \( N \) independent eigenvectors. This is the case if the characteristic polynomial, \( p(z) \), of the original scalar equation has distinct roots. If \( p(z) \) has multiple roots diagonalization may not be possible but \( C \) can always be reduced to Jordan (i.e. almost diagonal) form.

4.2 eigenvectors of the companion matrix \( C \)

Let \( p(z) \) be polynomial of degree \( N \) with complex coefficients

\[
p(z) = \sum_{k=1}^{N} a_k z^{N-k}
\]  

Then \( p \) generates the companion matrix

\[
C = \begin{pmatrix}
0 & 1 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
\frac{a_N}{a_0} & \frac{a_{N-1}}{a_0} & \frac{a_{N-2}}{a_0} & \frac{a_{N-3}}{a_0} & \cdots & \frac{a_2}{a_0} & \frac{a_1}{a_0}
\end{pmatrix}
\]  

The student might like to try the following exercise.

exercise

Let \( C \) be the \( N \times N \) companion matrix generated by the polynomial \( p(z) \) of degree \( N \). Then

(i)

\[
\det(zI - C) = \frac{1}{d_0} p(z)
\]  

(ii) The eigenvalues of \( C \) are the roots of the polynomial \( p(z) \).
Recall the vector valued function

\[ v: \mathbb{C} \ni z \mapsto v(z) = \begin{pmatrix} 1 \\ z \\ z^2 \\ z^3 \\ \vdots \\ z^{d-1} \\ \vdots \\ z^{N-2} \\ z^{N-1} \end{pmatrix} \] (4.4)

We will study the action of the companion matrix on the vector \( v(z) \).

\[
Cv(z) =
\begin{pmatrix}
0 & 1 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1 & 0 \\
\frac{a_N}{a_0} & -\frac{a_{N-1}}{a_0} & -\frac{a_{N-2}}{a_0} & -\frac{a_{N-3}}{a_0} & \cdots & -\frac{a_2}{a_0} & -\frac{a_1}{a_0}
\end{pmatrix}
\begin{pmatrix}
1 \\
z \\
z^2 \\
z^3 \\
\vdots \\
z^{d-1} \\
z^{N-2} \\
z^{N-1}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
z \\
z^2 \\
z^3 \\
z^4 \\
\vdots \\
z^i \\
z^{i+1} \\
z^{i+2} \\
z^{i+3} \\
\vdots \\
z^{N-1} \\
\frac{-1}{a_0}[p(z)-a_0z^N] \\
\end{pmatrix}
= \begin{pmatrix}
z \\
z^2 \\
z^3 \\
z^4 \\
\vdots \\
z^i \\
z^{i+1} \\
z^{i+2} \\
z^{i+3} \\
\vdots \\
z^{N-1} \\
\frac{-1}{a_0}[p(z)-a_0z^N] \\
\end{pmatrix}
= v(z) - \frac{p(z)}{a_0} e_N
\]

To summarize

\[ Cv(z) = zv(z) - \frac{p(z)}{a_0} v(z) \] (4.5)
If \( p(z_j) = 0 \) the next theorem follows immediately.

**Theorem 8** Let \( C \) be the companion matrix of the polynomial \( p(z) \). Let \( z \) be a root of \( p \). Then

\[
Cv(z) = zv(z)
\]

i.e. \( v(z) \) is an eigenvector of \( C \) with associated eigenvalue \( z \).

### 4.2.1 the case of no multiplicities

Assume that \( p(z) \) has \( N \) distinct roots, \( z_j, 1 \leq j \leq N \). Each root is an eigenvalue of the companion matrix \( C \).

\[
Cv(z_j) = z_jv(z_j), \quad 1 \leq j \leq N
\]

These \( N \) statements can be expressed in just one matrix statement (each eigenvector being a column of the Vandermonde matrix \( A = A(z_1, z_2, z_3, \cdots, z_j, \cdots, z_N) \)).

\[
CA(z_1, z_2, z_3, \cdots, z_j, \cdots, z_N) = A(z_1, z_2, z_3, \cdots, z_j, \cdots, z_N)D(z_1, z_2, z_3, \cdots, z_j, \cdots, z_N)
\]

\[
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & 1 \\
(z_1)^{N-1} & (z_2)^{N-1} & (z_3)^{N-1} & \cdots & (z_j)^{N-1} & (z_N)^{N-1} \\
(z_1)^{N-2} & (z_2)^{N-2} & (z_3)^{N-2} & \cdots & (z_j)^{N-2} & (z_N)^{N-2} \\
(z_1)^{N-3} & (z_2)^{N-3} & (z_3)^{N-3} & \cdots & (z_j)^{N-3} & (z_N)^{N-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(z_1) & z_2 & z_3 & \cdots & z_j & z_N
\end{pmatrix}
\]

\[
= \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & 1 \\
(z_1)^{N-1} & (z_2)^{N-1} & (z_3)^{N-1} & \cdots & (z_j)^{N-1} & (z_N)^{N-1} \\
(z_1)^{N-2} & (z_2)^{N-2} & (z_3)^{N-2} & \cdots & (z_j)^{N-2} & (z_N)^{N-2} \\
(z_1)^{N-3} & (z_2)^{N-3} & (z_3)^{N-3} & \cdots & (z_j)^{N-3} & (z_N)^{N-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(z_1) & z_2 & z_3 & \cdots & z_j & z_N
\end{pmatrix}
\]

We write for short

\[
C = A^{-1}DA \quad \Leftrightarrow A^{-1}CA = D
\]
Here $D$ is the diagonal matrix with the roots $\{z_j \mid 1 \leq j \leq N\}$ along the diagonal, all other entries being zero.

## 4.2.2 the case with multiplicities

Assume that

$$p(z) = a_0 \prod_{j=1}^{r} (z-z_j)^{m_j}$$

i.e. $p$ has $r$ roots only, $\{z_j \mid 1 \leq j \leq r\}$ where the $j$-th root $z_j$ has multiplicity $m_j$, $1 \leq j \leq r$. Proceeding as in §4.2.1 we obtain only $r < N$ eigenvectors of the companion matrix $C$, not enough for diagonalization. In fact diagonalization of $C$ is impossible if $p$ has a multiple root; consider

**exercise** Let $p(z) = (z-a)^2$.

(i) The companion matrix is

$$C = \begin{pmatrix} 0 & 1 \\ -a^2 & 1 \end{pmatrix}$$

(ii) $C$ possesses only one eigenvector; find it and prove there is no other (up to scalar multiplication).

(iii) Prove that $C$ is undiagonizable, i.e. there is no $2 \times 2$ matrix $A$ such that $A^{-1}CA = D$, a $2 \times 2$ diagonal matrix.

We will apply the companion matrix $C$ to each column of the confluent Vandermond matrix. Let $\alpha > 0$, then

$$C \frac{1}{\alpha!} D_{\alpha}^z v(z)$$

$$= \frac{1}{\alpha!} D_{\alpha}^z C v(z) , \quad \text{constant matrix } C \text{ commutes with } D_{\alpha}$$

$$= \frac{1}{\alpha!} D_{\alpha}^z \left[ z v(z) - \frac{p(z)}{a_0} e_N \right] , \quad \text{by } \S 4.2$$

$$= \frac{1}{\alpha!} D_{\alpha}^z (z v(z)) - \frac{p^{(\alpha)}(z)}{a_0} e_N$$

$$= \frac{1}{\alpha!} \left[ \sum_{k=0}^{\alpha} \binom{\alpha}{k} D_{\alpha}^z D_{\alpha}^{z-1} v(z) - \frac{p^{(\alpha)}(z)}{a_0} e_N \right]$$

$$= \frac{1}{\alpha!} \left[ z D_{\alpha}^z v(z) + \alpha D_{\alpha}^{z-1} v(z) + 0 + 0 \cdots + 0 \right] - \frac{p^{(\alpha)}(z)}{a_0^2} e_N$$

$$= \frac{1}{\alpha!} \left[ \frac{1}{\alpha - 1} D_{\alpha}^{z-1} v(z) - \frac{p^{(\alpha)}(z)}{a_0^2} e_N \right]$$

Next recall $z_j$, $1 \leq j \leq r$, each a root of $p(z)$ with multiplicity $m_j \geq 1$. Thus $(z-z_j)^{m_j}$ is a factor of $p(z)$, $1 \leq j \leq r$ and so

$$p^{(\alpha)}(z_j) = D_{\alpha}^z p(z_j) = 0 , \quad 0 \leq \alpha \leq m_j - 1 , \quad 1 \leq j \leq r$$

Thus

$$C v(z_j) = z_j v(z_j) , \quad 1 \leq j \leq r \quad (4.10)$$

and

$$C \left( \frac{1}{\alpha!} D_{\alpha}^z v(z_j) \right) = z_j \frac{1}{\alpha!} D_{\alpha}^z v(z_j) + \frac{1}{(\alpha - 1)!} D_{\alpha-1}^{z} v(z_j) , \quad 1 \leq j \leq r , \quad 1 \leq \alpha \leq m_j - 1 \quad (4.11)$$
The last two equations tell us that some columns of the confluent Vandermond matrix are eigenvectors of $C$ but some are not: but even these are almost eigenvectors This list of $N$ statements above can be encapsulated in one matrix equation which works because the Jordan form matrix contains in, its superdiagonal line, terms which express the failure to be eigenvectors of the confluent Vandermond columns.

$$CA(z_1, z_2, \ldots, z_j, \ldots, z_m) = A(z_1, z_2, \ldots, z_j, \ldots, z_m)J(z_1, z_2, \ldots, z_j, \ldots, z_m)$$ \hspace{1cm} \text{(4.12)}

which can be written for short as

$$A^{-1}CA = J \iff C = AJA^{-1}$$ \hspace{1cm} \text{(4.13)}

where

$$A = A(z_1, z_2, \ldots, z_j, \ldots, z_m)$$ \hspace{1cm} \text{(4.14)}

is the confluent Vandermond matrix and

$$J = J(z_1, z_2, \ldots, z_j, \ldots, z_m) = \begin{pmatrix}
J(m_1, z_1) & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & J(m_2, z_2) & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & J(m_3, z_3) & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & J(m_4, z_4) & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & J(m_k, z_k) & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\end{pmatrix}$$

called Jordan Form of $C$, where for any integer $m \geq 1$ and $z \in \mathbb{C}$, $J(m, z)$, called a basic Jordan block, is the following $m \times m$ matrix and each zero in the latter matrix denotes a block of zeros of appropriate size.

$$J(m, z) = \begin{pmatrix}
z & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & z & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & z & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & z & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & z & 1 & \cdots & 0 & \cdots & z \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & z & \cdots & 0 & \cdots & 1 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & z & 1 & \cdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & z \\
\end{pmatrix}$$

\textbf{4.3 examples of solution of IHDIEqns}

Given a SIHDIE

(i) Convert to a IHVDIE involving $C$ the companion matrix

(ii) Find $A$ (confluent Vandermond) and $J$ (Jordan Form) so that $C = AJA^{-1}$; i.e. quasi diagonalize $C$, reducing it by similarity transform to Jordan Form.

(iii) Solve using the formula (with matrix vector convolution)

$$x(t) = \exp(tC)x(0) + \exp(tC)*h(t)$$

(iv) Obtain the scalar $x(t)$ as the component $x^1(t)$

For the present, see lecture notes and past exams with their model solutions for actual examples.
4.4 examples of solution of IHDcEqns

Given a SIHDcE

(i) Convert to a IHVDcE involving $C$ the companion matrix

(ii) Find $A$ (confluent Vandermond) and $J$ (Jordan Form) so that $C = AJA^{-1}$; i.e. quasi diagonalize $C$, reducing it by similarity transform to Jordan Form.

(iii) Solve using the formula (with matrix vector convolution)

$$u_n = C^n u_0 + C^n * b_n$$

(iv) Obtain the scalar $u_n$ as the component $(u_n)^1$

For the present, see lecture notes and past exams with their model solutions for actual examples.