CHAPTER 6

Application of Jordan Canonical Form

Notations

• \( \mathbb{R} \) is the set of real numbers.
• \( \mathbb{C} \) is the set of complex numbers.
• \( \mathbb{Q} \) is the set of rational numbers.
• \( \mathbb{Z} \) is the set of integers.
• \( \mathbb{N} \) is the set of non-negative integers.
• \( \mathbb{Z}^+ \) is the set of positive integers.
• \( \text{Re}(z) \), \( \text{Im}(z) \), \( \overline{z} \) and \( |z| \) are the real part, imaginary part, conjugate and modulus of a complex number \( z \).
• \( M_{m \times n}(F) \) is the set of \( m \times n \) matrices with entries in \( F \).
• \( F(x) \) is the set of polynomials in \( x \) with coefficients in \( x \).
• \( \text{Row}(A) \), \( \text{Col}(A) \) and \( \text{Nul}(A) \) are the row, column and null spaces of a matrix \( A \), respectively.
• \( \text{rank}(A) \) is the rank of a matrix \( A \).
• \( I_n = [e_1 \ e_2 \ldots \ e_n] \) is the \( n \times n \) identity matrix, where \( e_1, e_2, \ldots, e_n \) are the column vectors of \( I_n \).
• \( K(T) \), \( R(T) \) and \( \text{rank}(T) \) are the kernel, range and rank of a linear transformation \( T \).
• \( \text{gcd}(f_1(x), f_2(x), \ldots, f_m(x)) \) is the greatest common divisor of polynomials \( f_1(x), f_2(x), \ldots, f_m(x) \).
• \( \text{lcm}(f_1(x), f_2(x), \ldots, f_m(x)) \) is the least common multiple of polynomials \( f_1(x), f_2(x), \ldots, f_m(x) \).

1. Computation of \( A^n \)

In many applications, we need to compute the \( n \)-th power of a given square matrix \( A \) (cf. \([P]\)). Here by computing \( A^n \), we mean to given a closed formula for \( A^n \) that depends only on \( n \) for a given \( A \). Our main tools are Jordan Canonical Form and characteristic/minimal polynomials.

There are essentially two approaches to the problem. The first method uses directly the JCF of \( A \). For every complex square matrix \( A \). There exists an invertible matrix \( P \) such that \( P^{-1}AP = J \) is a Jordan matrix. Then \( A = PJP^{-1} \) and

\[
A^n = PJ^nP^{-1}.
\]
6. APPLICATION OF JORDAN CANONICAL FORM

It suffices to figure out the \( n \)-th power of a Jordan matrix, which comes down to computing the \( n \)-th power of the Jordan block \( J_{\lambda,m} \):

\[
J_{\lambda,m}^n = (\lambda I + J_{0,m})^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k} J_{0,m}^k
\]

(6.2)

\[
= \sum_{k=0}^{m-1} \binom{n}{k} \lambda^{n-k} J_{0,m}^k
\]

\[
= \binom{n}{0} \lambda^n I + \binom{n}{1} \lambda^{n-1} J_{0,m} + \ldots + \binom{n}{m-1} \lambda^{n-m+1} J_{0,m}^{m-1}
\]

by Binomial Theorem, since \( J_{0,m}^k = 0 \) for \( k \geq m \).

If we use the notation \( E_{ij} \) for the \( m \times m \) matrix with 1 at \( i \)-th row and \( j \)-th column and 0 everywhere else, then

\[
J_{0,m} = E_{12} + E_{23} + \ldots + E_{m-1,m} = \sum_{i=1}^{m-1} E_{i,i+1}
\]

\[
J_{0,m}^k = E_{1,k+1} + E_{2,k+2} + \ldots + E_{m-k,m} = \sum_{i=1}^{m-k} E_{i,i+k}
\]

(6.3)

\[
J_{\lambda,m}^n = \sum_{k=0}^{m-1} \binom{n}{k} \lambda^{n-k} J_{0,m}^k = \sum_{k=0}^{m-1} \binom{n}{k} \lambda^{n-k} \left( \sum_{i=1}^{m-k} E_{i,i+k} \right)
\]

\[
= \sum_{k=0}^{m-1} \sum_{i=1}^{m-k} \binom{n}{k} \lambda^{n-k} E_{i,i+k} = \sum_{1 \leq i \leq j \leq m} \binom{n}{j-i} \lambda^{n+i-j} E_{ij}.
\]

So \( J_{\lambda,m}^n \) is an upper triangular matrix whose entry at \( i \)-th row and \( j \)-th column is \( \binom{n}{j-i} \lambda^{n+i-j} \) for \( i \leq j \). That is,

\[
J_{\lambda,m}^n = \begin{bmatrix}
\binom{n}{0} \lambda^n & \binom{n}{1} \lambda^{n-1} & (\frac{n}{2}) \lambda^{n-2} & \ldots & (\frac{n}{m-1}) \lambda^{n-m+1} \\
\binom{n}{0} \lambda^n & \binom{n}{1} \lambda^{n-1} & (\frac{n}{2}) \lambda^{n-2} & \ldots & (\frac{n}{m-2}) \lambda^{n-m+2} \\
& \ddots & \ddots & \ddots & \ddots \\
\binom{n}{0} \lambda^n & \binom{n}{1} \lambda^{n-1} & (\frac{n}{2}) \lambda^{n-2} & \ldots & \binom{n}{0} \lambda^n
\end{bmatrix}
\]

(6.4)

For example,

\[
J_{\lambda,2}^n = \begin{bmatrix}
\lambda & 1^n \\
\lambda & \lambda
\end{bmatrix} = \begin{bmatrix}
\binom{n}{0} \lambda^n & \binom{n}{1} \lambda^{n-1} \\
\binom{n}{0} \lambda^n & \binom{n}{1} \lambda^{n-1}
\end{bmatrix} = \begin{bmatrix}
\lambda^n & n\lambda^{n-1} \\
\lambda^n & \lambda^n
\end{bmatrix}.
\]

**Example 6.1.** Let us compute \( A^n \) for

\[
A = \begin{bmatrix}
8 & -4 \\
9 & -4
\end{bmatrix}.
\]

The characteristic polynomial of \( A \) is \((x - 2)^2\) and

\[
\{ \dim \text{Nul}(A - 2I)^k : k = 0, 1, 2, \ldots \} = \{0, 1, 2, 2, \ldots \}.
\]
So the Young tableau of $A$ is given in Figure 7. It suffices to choose a vector $v \not\in \text{Nul}(A - 2I)$ and then

$$P^{-1}AP = J_{2,2} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}$$

for $P = [(A - 2I)v \quad v]$.

We may choose $v = e_1$. Then

$$P = \begin{bmatrix} 6 & 1 \\ 9 & 0 \end{bmatrix}$$

$$A = PJ_{2,2}P^{-1}$$

$$A^n = PJ_{2,2}^nP^{-1} = \begin{bmatrix} 6 & 1 \\ 9 & 0 \end{bmatrix}^n \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}^{-1}$$

$$= -\frac{1}{9} \begin{bmatrix} 6 & 1 \\ 9 & 0 \end{bmatrix}^n \begin{bmatrix} 2^n & n(2^{n-1}) \\ 2^n & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -9 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} (3n + 1)2^n & -n(2^{n+1}) \\ 9n(2^{n-1}) & (1 - 3n)2^n \end{bmatrix}.$$

An equivalent way to compute $A^n$ is to compute $A^n v$ by decomposing $v$ to a sum of generalized eigenvectors of $A$.

The column vectors of the invertible matrix $P$ in (6.1) form a basis $B$ of $\mathbb{C}^n$ which is a disjoint union

$$B = B_1 \sqcup B_2 \sqcup \ldots \sqcup B_l$$

of cycles of generalized eigenvectors. Suppose that each $B_i$ is generated by the vector $v_i$ corresponding to eigenvalue $\lambda_i$ and $|B_i| = m_i$. To compute $A^n v$, it suffices to write $v$ as a linear combination of $B$:

$$v = \sum_{i=1}^l \sum_{j=0}^\infty c_{ij} (A - \lambda_i I)^j v_i$$

$$(6.11)$$

$$\Rightarrow A^n v = \sum_{i=1}^l A^n \sum_{j=0}^\infty c_{ij} (A - \lambda_i I)^j v_i$$

where we may take $j$ from 0 to $\infty$ for simplicity since $(A - \lambda_i I)^j v_i = 0$ for $j \geq m_i$.

The term $A^n \sum(A - \lambda_i I)^j v_i$ can be computed by expanding $A^n$ as

$$A^n = (\lambda_i I + (A - \lambda_i I))^n = \sum_{k=0}^n \binom{n}{k} \lambda_i^{n-k} (A - \lambda_i I)^k$$

$$(6.12)$$

$$= \sum_{k=0}^\infty \binom{n}{k} \lambda_i^{n-k} (A - \lambda_i I)^k.$$
Therefore,

\[
A^n v = \sum_{i=1}^l A^n \sum_{j=0}^\infty c_{ij} (A - \lambda_i I)^j v_i
\]

(6.13)

\[
= \sum_{i=1}^l \left( \sum_{k=0}^{\infty} \left( \begin{array}{c} n \\ k \end{array} \right) \lambda_i^{n-k} (A - \lambda_i I)^k \right) \left( \sum_{j=0}^\infty c_{ij} (A - \lambda_i I)^j v_i \right)
\]

\[
= \sum_{i=1}^l \sum_{j,k \geq 0} c_{ij} \lambda_i^{n-k} \left( \begin{array}{c} n \\ k \end{array} \right) (A - \lambda_i I)^{j+k} v_i.
\]

Letting \( j + k = a \) and \( k = a - j \), we can rewrite (6.13) as

\[
A^n v = \sum_{i=1}^l \sum_{j,k \geq 0} c_{ij} \lambda_i^{n-k} \left( \begin{array}{c} n \\ k \end{array} \right) (A - \lambda_i I)^{j+k} v_i
\]

(6.14)

\[
= \sum_{i=1}^l \sum_{a=0}^{\infty} \sum_{j=0}^a c_{ij} \lambda_i^{n+j-a} \left( \begin{array}{c} n \\ a-j \end{array} \right) (A - \lambda_i I)^a v_i
\]

\[
= \sum_{i=1}^l \sum_{a=0}^{m_i-1} \sum_{j=0}^a c_{ij} \lambda_i^{n+j-a} \left( \begin{array}{c} n \\ a-j \end{array} \right) (A - \lambda_i I)^a v_i
\]

since \((A - \lambda_i I)^a v_i = 0\) for \( a \geq m_i \).

So we may use (6.14) to compute \( A^n v \) for all \( v \). To compute \( A^n \) itself, we observe that

\[
A^n = A^n I = A^n [e_1 \quad e_2 \ldots \quad e_m] = [A^n e_1 \quad A^n e_2 \ldots \quad A^n e_m]
\]

(6.15)

and apply the above algorithm to each \( A^n e_i \).

**Example 6.2.** Let us redo Example 6.1 by computing \( A^n e_i \) via writing \( e_i \) as a linear combination of \( B = \{(A - 2I)v, v\} \), where we choose \( v = e_1 \). Then

\[
e_1 = e_1
\]

(6.16)

\[
e_2 = -\frac{2}{3} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{9} \begin{bmatrix} 6 \\ 9 \end{bmatrix} = -\frac{2}{3} e_1 + \frac{1}{9} (A - 2I)e_1
\]
and hence

\[ A^n e_1 = A^n e_1 = (2I + (A - 2I))^n e_1 \]

\[ = \sum_{k=0}^{n} \binom{n}{k} 2^{n-k} (A - 2I)^k e_1 \]

\[ = \sum_{k=0}^{1} \binom{n}{k} 2^{n-k} (A - 2I)^k e_1 = 2^n e_1 + n(2^{n-1})(A - 2I)e_1 \]

\[ = 2^n \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] + n(2^{n-1}) \left[ \begin{array}{c} 6 \\ 9 \end{array} \right] = \left[ \begin{array}{c} (3n + 1)2^n \\ 9n(2^{n-1}) \end{array} \right] \]

\[ A^n e_2 = A^n \left( -\frac{2}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right) \]

\[ = (2I + (A - 2I))^n \left( -\frac{2}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right) \]

\[ = \left( \sum_{k=0}^{n} \binom{n}{k} 2^{n-k} (A - 2I)^k \right) \left( -\frac{2}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right) \]

\[ = (2^n I + n(2^{n-1})(A - 2I)) \left( -\frac{2}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right) \]

\[ = -\frac{2^{n+1}}{3} e_1 + \left( \frac{2^n}{9} - \frac{n(2^n)}{3} \right) (A - 2I)e_1 \]

\[ = -\frac{2^{n+1}}{3} \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] + \left( \frac{2^n}{9} - \frac{n(2^n)}{3} \right) \left[ \begin{array}{c} 6 \\ 9 \end{array} \right] = \left[ \begin{array}{c} -n(2^{n+1}) \\ (1 - 3n)2^n \end{array} \right] . \]

Therefore,

\[ A^n = \left[ \begin{array}{ll} A^n e_1 & A^n e_2 \end{array} \right] = \left[ \begin{array}{cc} (3n + 1)2^n & -n(2^{n+1}) \\ 9n(2^{n-1}) & (1 - 3n)2^n \end{array} \right] . \]

The main complexity of the above algorithm of computing \( A^n \) via JCF lies in the finding of the invertible matrix \( P \), which is a time-consuming process. Our second method sidesteps the whole issue of \( P \), but it introduces its own complexity.

The main idea is that if we have a nonzero polynomial \( f(x) \) such that \( f(A) = 0 \), we can divide \( x^n \) by \( f(x) \) to obtain

\[ x^n = q(x)f(x) + r(x) \]

where \( q(x) \) and \( r(x) \) is the quotient and remainder of the division of \( x^n \) by \( f(x) \) such that \( \deg r(x) < \deg f(x) \); then

\[ A^n = q(A)f(A) + r(A) = r(A) \]

since \( f(A) = 0 \). Of course, \( q(x) \) is irrelevant here. The key of the above approach is the determination of the polynomial \( r(x) \).

We may take \( f(x) \) to be any nonzero polynomial satisfying \( f(A) = 0 \), i.e., any nonzero polynomial that is divisible by the minimal polynomial of
A. But the lesser the degree of \( f(x) \), the better the above algorithm works. If the minimal polynomial of \( A \) is know, we take \( f(x) \) to be the minimal polynomial. Otherwise, we usually take \( f(x) \) to be the characteristic polynomial of \( A \) by Cayley-Hamilton.

To determine \( r(x) \), we write (6.19) as

\[
(6.21) \quad \frac{x^n}{f(x)} = \frac{q(x)}{f(x)} + \frac{r(x)}{f(x)}.
\]

Let \( f(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \ldots (x - \lambda_m)^{d_m} \). We may write \( r(x)/f(x) \) as partial fractions:

\[
(6.22) \quad \frac{x^n}{f(x)} = \frac{q(x)}{f(x)} = q(x) + \sum_{i=1}^{m} \sum_{j=1}^{d_i} \frac{c_{ij}}{(x - \lambda_i)^j}
\]

for some numbers \( c_{ij} \). To determine \( c_{ij} \) for each \( i \), we multiply both sides of (6.22) by \( (x - \lambda_i)^{d_i} \) and take derivatives at \( \lambda_i \) of order \( k = 0, 1, \ldots, d_i - 1 \) in sequence:

\[
(6.23) \quad \left. \left( \frac{x^n(x - \lambda_i)^{d_i}}{f(x)} \right)^{(k)} \right|_{\lambda_i} = \left. \left( \sum_{j=1}^{d_i} c_{ij}(x - \lambda_i)^{d_i-j} \right)^{(k)} \right|_{\lambda_i}
\]

\[ + \left. \left( h_i(x)(x - \lambda_i)^{d_i} \right)^{(k)} \right|_{\lambda_i} \]

where

\[
(6.24) \quad h_i(x) = q(x) + \sum_{l \neq i}^{m} \sum_{j=1}^{d_l} \frac{c_{lj}}{(x - \lambda_l)^j}.
\]

Since \( h_i(x) \) is \( C^\infty \) at \( \lambda_i \), it is easy to see that

\[
(6.25) \quad \left. \left( h_i(x)(x - \lambda_i)^{d_i} \right)^{(k)} \right|_{\lambda_i} = 0
\]

for all \( k < d_i \). So

\[
(6.26) \quad \left. \left( \frac{x^n(x - \lambda_i)^{d_i}}{f(x)} \right)^{(k)} \right|_{\lambda_i} = \left. \left( \sum_{j=1}^{d_i} c_{ij}(x - \lambda_i)^{d_i-j} \right)^{(k)} \right|_{\lambda_i}
\]

\[ = (k!)c_{i,d_i-k} \]

for \( k = 0, 1, \ldots, d_i - 1 \) and hence

\[
(6.27) \quad c_{ij} = \frac{1}{(d_i - j)!} \left. \left( \frac{x^n(x - \lambda_i)^{d_i}}{f(x)} \right)^{(d_i-j)} \right|_{\lambda_i}.
\]

In complex analysis, \( c_{ij} \) is the residue of the meromorphic function

\[
(6.28) \quad \frac{z^n(z - \lambda_i)^{j-1}}{f(z)}
\]
at $z = \lambda_i$.

Example 6.3. For example, suppose that $f(x) = (x - 1)^2(x - 2)^3$ and we write

$$
\frac{x^n}{(x - 1)^2(x - 2)^3} = q(x) + \frac{c_{11}}{x - 1} + \frac{c_{12}}{(x - 1)^2} + \frac{c_{21}}{x - 2} + \frac{c_{22}}{(x - 2)^2} + \frac{c_{23}}{(x - 2)^3}.
$$

(6.29)

To determine $c_{11}$ and $c_{12}$, we multiply both sides of (6.29) by $(x - 1)^2$ and take derivatives at 1 of order $k = 0, 1$:

$$
\left( \frac{x^n}{(x - 2)^3} \right)^{(k)} \bigg|_{x=1} = (c_{11}(x - 1) + c_{12})^{(k)} \bigg|_{x=1}
$$

(6.30)

$$
+ (h_1(x)(x - 1)^2)^{(k)} \bigg|_{x=1}
$$

where

$$
h_1(x) = q(x) + \frac{c_{21}}{x - 2} + \frac{c_{22}}{(x - 2)^2} + \frac{c_{23}}{(x - 2)^3}.
$$

(6.31)

So for $k = 0, 1$, we obtain

$$
c_{12} = \left( \frac{x^n}{(x - 2)^3} \right)^{(1)} \bigg|_{x=1} = -1
$$

(6.32)

$$
c_{11} = \left( \frac{x^n}{(x - 2)^3} \right)^{(2)} \bigg|_{x=1} = -n - 3.
$$

To determine $c_{21}, c_{22}$ and $c_{23}$, we multiply both sides of (6.29) by $(x - 2)^3$ and take derivatives at 2 of order $k = 0, 1, 2$: where

$$
\left( \frac{x^n}{(x - 1)^2} \right)^{(k)} \bigg|_{x=2} = (c_{21}(x - 2)^2 + c_{22}(x - 2) + c_{23})^{(k)} \bigg|_{x=2}
$$

(6.33)

$$
+ (h_2(x)(x - 2)^3)^{(k)} \bigg|_{x=2}
$$

where

$$
h_2(x) = q(x) + \frac{c_{11}}{x - 1} + \frac{c_{12}}{(x - 1)^2}.
$$

(6.34)

So for $k = 0, 1, 2$, we obtain

$$
c_{23} = \left( \frac{x^n}{(x - 1)^2} \right)_{x=2} = 2^n
$$

(6.35)

$$
c_{22} = \left( \frac{x^n}{(x - 1)^2} \right)'_{x=2} = (n - 4)2^{n-1}
$$

$$
c_{21} = \frac{1}{2} \left( \frac{x^n}{(x - 1)^2} \right)''_{x=2} = (n^2 - 9n + 24)2^{n-3}.
$$
Therefore,

\[
\frac{x^n}{(x-1)^2(x-2)^3} = q(x) + \frac{n+3}{x-1} - \frac{1}{(x-1)^2} \\
+ \frac{(n^2-9n+24)2^{n-3}}{x-2} + \frac{(n-4)2^{n-1}}{(x-2)^2} + \frac{2^n}{(x-2)^3}
\]

and hence

\[
x^n = q(x)f(x) - (n+3)(x-1)(x-2)^3 - (x-2)^3 \\
+ (n^2-9n+24)2^{n-3}(x-1)^2(x-2)^2 \\
+ (n-4)2^{n-1}(x-1)^2(x-2) + 2^n(x-1)^2.
\]

If \( f(A) = 0 \), then

\[
A^n = -(n+3)(A-I)(A-2I)^3 - (A-2I)^3 \\
+ (n^2-9n+24)2^{n-3}(A-I)^2(A-2I)^2 \\
+ (n-4)2^{n-1}(A-I)^2(A-2I) + 2^n(A-I)^2.
\]

**Example 6.4.** Let us redo Example 6.1 using the characteristic polynomial of \( A \). Let \( f(x) = \text{det}(xI - A) = (x-2)^2 \) and let

\[
\frac{x^n}{(x-2)^2} = q(x) + \frac{c_1}{x-2} + \frac{c_2}{(x-2)^2}
\]

for some numbers \( c_1 \) and \( c_2 \). Then

\[
c_1 = (x^n)' \bigg|_{x=2} = n(2^{n-1}) \quad \text{and} \quad c_2 = x^n \bigg|_{x=2} = 2^n.
\]

Therefore,

\[
x^n = q(x)f(x) + n(2^{n-1})(x-2) + 2^n \\
A^n = n(2^{n-1})(A-2I) + 2^nI \\
= n(2^{n-1})A + (1-n)2^nI
\]

\[
= n(2^{n-1}) \begin{bmatrix} 8 & -4 \\ 9 & -4 \end{bmatrix} + (1-n)2^n \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\]

\[
= \begin{bmatrix} (3n+1)2^n & -n(2^{n+1}) \\ 9n(2^{n-1}) & (1-3n)2^n \end{bmatrix}.
\]

**Example 6.5.** Let us compute \( A^n \) for

\[
A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.
\]

Let \( f(x) = \text{det}(xI - A) = (x-1)^2(x-4) \) and let

\[
\frac{x^n}{(x-1)^2(x-4)} = q(x) + \frac{a_1}{x-1} + \frac{a_2}{(x-1)^2} + \frac{b}{x-4}
\]
1. Computation of $A^n$

for some numbers $a_1, a_2$ and $b$. Then

\begin{equation}
\begin{aligned}
a_1 &= \left. \frac{x^n}{x-4} \right|_{x=1} = -\frac{n}{3} - \frac{1}{9}, \\
b &= \left. \frac{x^n}{(x-1)^2} \right|_{x=4} = \frac{4^n}{9}.
\end{aligned}
\end{equation}

Therefore,

\begin{align*}
\frac{x^n}{(x-1)(x-4)} &= q(x)f(x) - \left( \frac{n}{3} + \frac{1}{9} \right) (x-1)(x-4) - \frac{1}{3}(x-4) + \frac{4^n}{9}(x-1)^2 \\
A^n &= - \left( \frac{n}{3} + \frac{1}{9} \right) (A-I)(A-4I) - \frac{1}{3}(A-4I) + \frac{4^n}{9}(A-I)^2 \\
\text{(6.45)} &= -\frac{1}{3} \begin{bmatrix}
-2 & 1 & 1 \\
1 & -2 & 1 \\
1 & 1 & -2
\end{bmatrix} + \frac{4^n}{9} \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{bmatrix}^2 \\
&= \frac{1}{3} \begin{bmatrix}
4^n + 2 & 4^n - 1 & 4^n - 1 \\
4^n - 1 & 4^n + 2 & 4^n - 1 \\
4^n - 1 & 4^n - 1 & 4^n + 2
\end{bmatrix}.
\end{align*}

Actually, $A$ is diagonalizable since

\begin{equation}
\text{dim Nul}(A-I) + \text{dim Nul}(A-4I) = 2 + 1 = 3
\end{equation}

and the minimal polynomial of $A$ is $(x-1)(x-4)$. So we may choose $f(x) = (x-1)(x-4)$ and the above computation can be simplified:

\begin{align*}
\frac{x^n}{(x-1)(x-4)} &= q(x) + \frac{a}{x-1} + \frac{b}{x-4} \\
\text{(6.47)} &= q(x) - \frac{1}{3(x-1)} + \frac{4^n}{3(x-4)}
\end{align*}

and hence

\begin{align*}
x^n &= q(x)f(x) - \frac{1}{3}(x-4) + \frac{4^n}{3}(x-1) \\
A^n &= -\frac{1}{3}(A-4I) + \frac{4^n}{3}(A-I) \\
\text{(6.48)} &= -\frac{1}{3} \begin{bmatrix}
-2 & 1 & 1 \\
1 & -2 & 1 \\
1 & 1 & -2
\end{bmatrix} + \frac{4^n}{3} \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{bmatrix} \\
&= \frac{1}{3} \begin{bmatrix}
4^n + 2 & 4^n - 1 & 4^n - 1 \\
4^n - 1 & 4^n + 2 & 4^n - 1 \\
4^n - 1 & 4^n - 1 & 4^n + 2
\end{bmatrix}.
\end{align*}

As an application of computation of $A^n$, we study linear recurrence.

A linear recurrence is a sequence \{a_0, a_1, \ldots, \} satisfying the condition

\begin{equation}
a_{n+1} = c_1a_{n+1-1} + c_2a_{n+1-2} + \ldots + c_0a_n + b
\end{equation}
for all \( n \geq 1 \), where \( l, c_1, c_2, \ldots, c_l, b \) are constants. If \( b = 0 \), it is homogenous. If \( b \neq 0 \), it is non-homogenous.

For example, a geometric sequence/progression is a homogeneous linear recurrence given by

\[
(6.50) \quad a_{n+1} = ra_n
\]

and an arithmetic sequence/progression is a non-homogeneous linear recurrence given by

\[
(6.51) \quad a_{n+1} = a_n + b.
\]

The most famous linear recurrence (other than geometric/arithmetic) is the Fibonacci sequence

\[
(6.52) \quad a_{n+2} = a_{n+1} + a_n.
\]

Solving a linear recurrence involves finding a closed formula for \( a_n \).

To solve \((6.49)\) for \( b = 0 \), we write

\[
(6.53) \quad \begin{cases} 
    a_{n+l} = c_1 a_{n+l-1} + c_2 a_{n+l-2} + \ldots + c_l a_n \\
    a_{n+l-1} = a_{n+l-1} \\
    a_{n+l-2} = a_{n+l-2} \\
    \vdots = \vdots \\
    a_{n+1} = a_{n+1} 
\end{cases}
\]

Let us put both sides of \((6.53)\) in the matrix form:

\[
(6.54) \quad \begin{bmatrix} a_{n+l} \\ a_{n+l-1} \\ \vdots \\ a_{n+1} \end{bmatrix}_{u_{n+1}} = \begin{bmatrix} c_1 & c_2 & \ldots & c_{l-1} & c_l \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}_A \begin{bmatrix} a_{n+l-1} \\ a_{n+l-2} \\ \vdots \\ a_n \end{bmatrix}_{u_n}
\]

This turns the linear recurrence \((6.49)\) into a “geometric” sequence \( \{u_n\} \) of vectors

\[
(6.55) \quad u_{n+1} = Au_n \text{ for } u_n = \begin{bmatrix} a_{n+l-1} \\ a_{n+l-2} \\ \vdots \\ a_n \end{bmatrix}
\]

in \( \mathbb{C}^l \). Obviously,

\[
(6.56) \quad u_n = Au_{n-1} = A^2 u_{n-2} = \ldots = A^n u_0 \text{ for } u_0 = \begin{bmatrix} a_{l-1} \\ a_{l-2} \\ \vdots \\ a_0 \end{bmatrix}
\]

So it comes down to the computation of \( A^n \) with \( A \) given in \((6.54)\).
When \( b \neq 0 \), it suffices to add one more component to the vectors \( \mathbf{u}_n \):

\[
\begin{bmatrix}
    a_{n+1} \\
    a_{n+1-l} \\
    \vdots \\
    a_{n+1-l-1} \\
    1
\end{bmatrix} =
\begin{bmatrix}
    c_1 & c_2 & \ldots & c_{l-1} & c_l & b \\
    1 & 0 & \ldots & 0 & 0 & 0 \\
    0 & 1 & \ldots & 0 & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
    0 & 0 & \ldots & 1 & 0 & 0 \\
    0 & 0 & \ldots & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    a_{n+l-1} \\
    a_{n+l-2} \\
    \vdots \\
    a_{n+1} \\
    1
\end{bmatrix}.
\]

Then

\[
\mathbf{u}_n = A^n \mathbf{u}_0 \quad \text{for} \quad \mathbf{u}_0 =
\begin{bmatrix}
    a_{l-1} \\
    a_{l-2} \\
    \vdots \\
    a_0 \\
    1
\end{bmatrix}.
\]

The characteristic polynomial of \( A \) for \( b = 0 \) can be computed by expanding the determinant of \( xI - A \) by the first row:

\[
\det(xI - A) = \det
\begin{bmatrix}
    x - c_1 & -c_2 & \ldots & -c_{l-1} & -c_l \\
    -1 & x & \ldots & 0 & 0 \\
    0 & -1 & \ldots & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \ldots & -1 & x
\end{bmatrix}
\]

\[
= (x - c_1) \det
\begin{bmatrix}
    x \\
    -1 & x \\
    \vdots & \vdots \\
    -1 & x
\end{bmatrix} - (-c_2) \det
\begin{bmatrix}
    -1 \\
    -1 & x \\
    \vdots & \vdots \\
    -1 & x
\end{bmatrix}
\]

\[
+ \cdots + (-1)^{l+2}(-c_l) \det
\begin{bmatrix}
    -1 & x \\
    -1 & x \\
    \vdots & \vdots \\
    -1 & x
\end{bmatrix}
\]

\[
= x^l - c_1 x^{l-1} - c_2 x^{l-2} - \cdots - c_l.
\]

Similarly, we can compute the characteristic polynomial of \( A \) for \( b \neq 0 \):

\[
\det(xI - A) = (x^l - c_1 x^{l-1} - c_2 x^{l-2} - \cdots - c_l)(x - 1)
\]

where \( A \) is given in (6.57).
Regardless of \( b = 0 \) or \( b \neq 0 \), we need to compute \( A^n \) for the corresponding \( A \). Note that we are only interested in the component \( a_n \) in \( u_n \) so we only need to know the \( l \)-th row of \( A^n \).

Let \( J = P^{-1}AP \) be the JCF of \( A \) for some invertible matrix \( P \). Then

\[
(6.61) \quad u_n = A^n u_0 = PJ^n P^{-1} u_0
\]

which is

\[
(6.62) \quad \begin{bmatrix}
    a_{n+l-1} \\
    a_{n+l-2} \\
    \vdots \\
    a_n
\end{bmatrix}
= PJ^n P^{-1}
\begin{bmatrix}
    a_{l-1} \\
    a_{l-2} \\
    \vdots \\
    a_0
\end{bmatrix}
\]

if \( b = 0 \) and

\[
(6.63) \quad \begin{bmatrix}
    a_{n+l-1} \\
    a_{n+l-2} \\
    \vdots \\
    a_n
\end{bmatrix}
= PJ^n P^{-1}
\begin{bmatrix}
    a_{l-1} \\
    a_{l-2} \\
    \vdots \\
    a_0
\end{bmatrix}
\]

if \( b \neq 0 \), where \( a_0, a_1, \ldots, a_{l-1} \) are the first \( l \) terms of the sequence which are usually given.

For every Jordan block \( J_{0,m} \), \( J_{0,m}^n \) is given by (6.4). We observe that the entries of \( J_{0,m}^n \) as functions of \( n \), are in the vector space of

\[
(6.64) \quad \text{Span} \left\{ \lambda^n, \binom{n}{1} \lambda^{n-1}, \ldots, \binom{n}{m-1} \lambda^{n-m+1} \right\}
= \text{Span} \left\{ \binom{n}{k} \lambda^{n-k} : k = 0, 1, 2, \ldots, m - 1 \right\}
\]

For each \( k \), \( \binom{n}{k} \) is a polynomial in \( n \) of degree \( k \), i.e.,

\[
(6.65) \quad \binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k!} \in \text{Span} \left\{ 1, n, n^2, \ldots, n^k \right\}.
\]

Therefore, the space (6.64) is the same as

\[
(6.66) \quad \text{Span} \left\{ \lambda^n, \binom{n}{1} \lambda^{n-1}, \ldots, \binom{n}{m-1} \lambda^{n-m+1} \right\}
= \text{Span} \left\{ \lambda^n, n\lambda^n, \ldots, n^{m-1}\lambda^n \right\}
= \text{Span} \left\{ n^k\lambda^n : k = 0, 1, 2, \ldots, m - 1 \right\}
\]

which contains all the entries of \( J_{\lambda,m}^n \).

Suppose that

\[
(6.67) \quad \det(xI - A) = (x - \lambda_1)^{d_1} (x - \lambda_2)^{d_2} \ldots (x - \lambda_m)^{d_m}
\]
with $\lambda_1, \lambda_2, \ldots, \lambda_m$ distinct. Every Jordan block of $J$ is in the form of $J_{\lambda_i,a}$ for some $a \leq d_i$. So the entries of $J^n$, as functions of $n$, are in the space

$$\text{Span} \left\{ \lambda_1^n, n\lambda_1^n, \ldots, n^{d_1-1}\lambda_1^n, \right. \right.$$  
$$\left. \lambda_2^n, n\lambda_2^n, \ldots, n^{d_2-1}\lambda_2^n, \right. \right.$$  
$$\left. \cdots \right.$$  
$$\left. \lambda_m^n, n\lambda_m^n, \ldots, n^{d_m-1}\lambda_m^n \right\}$$  

(6.68)

Clearly, the entries of $PJ^nP^{-1}$ are also in the above space since $P$ is a constant matrix. It follows that $a_n$ lies in (6.68), i.e.,

$$a_n = \sum_{i=1}^{m} \sum_{j=0}^{d_i-1} c_{ij} n^j \lambda_i^n$$  

(6.69)

for some $c_{ij} \in \mathbb{C}$. These $c_{ij}$ can be determined then by the values of $a_0, a_1, \ldots, a_l$.

In summary, we can solve a linear recurrence in the following steps:

1. Determine the characteristic polynomial of the recurrence given by (6.59) or 6.60.
2. Find all the roots of the characteristic polynomial as in (6.67).
3. Write $a_n$ as a linear combination of the functions

$$\lambda_1^n, n\lambda_1^n, \ldots, n^{d_1-1}\lambda_1^n,$$

$$\lambda_2^n, n\lambda_2^n, \ldots, n^{d_2-1}\lambda_2^n,$$

$$\cdots$$

$$\lambda_m^n, n\lambda_m^n, \ldots, n^{d_m-1}\lambda_m^n$$

as (6.69).
4. Determine the coefficients $c_{ij}$ of this linear combination (6.69) by setting $n = 0, 1, \ldots, l$.

Alternatively, we may compute $a_n$ by computing $A^n$ directly using its characteristic polynomial.

**Example 6.6.** Let us solve the Fibonacci sequence $\{a_n : n = 0, 1, \ldots\}$ given by $a_{n+2} = a_{n+1} + a_n$ for $n \geq 0$, $a_0 = 0$ and $a_1 = 1$.

We can write

$$\begin{cases} a_{n+2} = a_{n+1} + a_n \\ a_{n+1} = a_n \end{cases} \iff \begin{bmatrix} a_{n+2} \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix}.$$  

(6.71)

It follows that

$$\begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix}^n \begin{bmatrix} a_1 \\ a_0 \end{bmatrix}.$$  

(6.72)
The characteristic polynomial of the recurrence is
\[
\det(xI - A) = x^2 - x - 1 = (x - \lambda_1)(x - \lambda_2)
\]
for \( \lambda_1 = \frac{1 + \sqrt{5}}{2} \) and \( \lambda_2 = \frac{1 - \sqrt{5}}{2} \).

So \( a_n \), as a function of \( n \), lies in the space spanned by \( \lambda_1^n \) and \( \lambda_2^n \). Therefore,
\[
a_n = c_1 \lambda_1^n + c_2 \lambda_2^n = c_1 \left( \frac{1 + \sqrt{5}}{2} \right)^n + c_2 \left( \frac{1 - \sqrt{5}}{2} \right)^n
\]
for some constants \( c_1 \) and \( c_2 \). Setting \( n = 0, 1 \), we have
\[
\begin{cases}
    c_1 + c_2 = 0 \\
    \lambda_1 c_1 + \lambda_2 c_2 = 1
\end{cases} \Rightarrow \begin{cases}
    c_1 = \frac{1}{\sqrt{5}} \\
    c_2 = -\frac{1}{\sqrt{5}}
\end{cases}
\]
and hence
\[
a_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^n.
\]

Alternatively, we may compute \( A^n \) directly via its characteristic polynomial \( f(x) = x^2 - x - 1 \). That is, we write
\[
\frac{x^n}{f(x)} = q(x) + \frac{b_1}{x - \lambda_1} + \frac{b_2}{x - \lambda_2}
\]
for some polynomial \( q(x) \) and some constants \( b_1 \) and \( b_2 \). Then
\[
b_1 = \left. \frac{x^n}{x - \lambda_2} \right|_{x = \lambda_1} = \frac{\lambda_1^n}{\lambda_1 - \lambda_2} \quad \text{and} \quad b_2 = \left. \frac{x^n}{x - \lambda_1} \right|_{x = \lambda_2} = \frac{\lambda_2^n}{\lambda_2 - \lambda_1}.
\]

By (6.77), we obtain
\[
x^n = q(x)f(x) + b_1(x - \lambda_2) + b_2(x - \lambda_1)
\]
and hence
\[
A^n = q(A)f(A) + b_1(A - \lambda_2 I) + b_2(A - \lambda_1 I)
\]
\[
= b_1(A - \lambda_2 I) + b_2(A - \lambda_1 I)
\]
\[
= (b_1 + b_2)A - (b_1 \lambda_2 + b_2 \lambda_1)I
\]
\[
= (b_1 + b_2) \begin{bmatrix} * & * \\ 1 & 1 \end{bmatrix} - (b_1 \lambda_2 + b_2 \lambda_1) \begin{bmatrix} * & * \\ 1 & 1 \end{bmatrix}
\]
\[
= \begin{bmatrix} * & * \\ b_1 + b_2 & -b_1 \lambda_2 - b_2 \lambda_1 \end{bmatrix}.
\]
Therefore,
\[
\begin{bmatrix}
a_{n+1} \\ a_n
\end{bmatrix} = A^n \begin{bmatrix} a_1 \\ a_0
\end{bmatrix}
\]
(6.81)
\[
= \begin{bmatrix}
* & * \\
\beta_1 + \beta_2 & -\beta_1 \lambda_2 - \beta_2 \lambda_1
\end{bmatrix} \begin{bmatrix} 1 \\ 0
\end{bmatrix}
= \begin{bmatrix} * \\
\beta_1 + \beta_2
\end{bmatrix}.
\]
So \(a_n = b_1 + b_2\) with \(b_1\) and \(b_2\) given in (6.78), which also gives us (6.76).

Example 6.7. Let us solve the recurrence \(\{a_n : n = 0, 1, 2, \ldots\}\) given by
\[
a_{n+2} = 3a_{n+1} - 2a_n + 5 \quad \text{for } n \geq 2 \quad \text{and } a_0 = a_1 = 1.
\]
We write
\[
\begin{cases}
a_{n+2} = 3a_{n+1} - 2a_n + 5 \\
a_{n+1} = a_n \\
1 = 1
\end{cases} \iff \begin{bmatrix} a_{n+2} \\ a_{n+1} \\ 1
\end{bmatrix} = \begin{bmatrix} 3 & -2 & 5 \\ 1 \\ 1
\end{bmatrix} \begin{bmatrix} a_n \\ a_{n+1} \\ 1
\end{bmatrix}.
\]
It follows that
\[
\begin{bmatrix} a_{n+1} \\ a_n \\ 1
\end{bmatrix} = \begin{bmatrix} 3 & -2 & 5 \\ 1 & 1 \\ 1
\end{bmatrix}^n \begin{bmatrix} a_1 \\ a_0 \\ 1
\end{bmatrix}.
\]
(6.84)
The characteristic polynomial of the recurrence is
\[
\det(xI - A) = (x^3 - 3x + 2)(x - 1) = (x - 1)^2(x - 2).
\]
So \(a_n\), as a function of \(n\), lies in the space spanned by \(1^n\), \(n(1^n)\) and \(2^n\).
Therefore,
\[
a_n = c_1 + c_2 n + c_3 (2^n)
\]
for some constants \(c_1, c_2\) and \(c_3\). Setting \(n = 0, 1, 2\), we have
\[
\begin{cases}
c_1 + c_2 = 1 \\
c_1 + c_2 + 2c_3 = 1 \Rightarrow c_2 = -5 \\
c_1 + 2c_2 + 4c_3 = 6 \Rightarrow c_3 = 5
\end{cases}
\]
and hence
\[
a_n = -4 - 5n + 5(2^n).
\]
Alternatively, we may compute \(A^n\) directly via its characteristic polynomial \(f(x) = (x - 1)^2(x - 2)\). That is, we write
\[
\frac{x^n}{f(x)} = q(x) + \frac{b_{11}}{x - 1} + \frac{b_{12}}{(x - 1)^2} + \frac{b_{21}}{x - 2}
\]
(6.89)
for some polynomial \( q(x) \) and some constants \( b_{ij} \). Then
\[
\begin{align*}
    b_{11} &= \left. \frac{x^n}{x-2} \right|_1 = -n - 1, \quad b_{12} = \left. \frac{x^n}{x-2} \right|_1 = -1 \quad \text{and} \\
    b_{21} &= \left. \frac{x^n}{(x-1)^2} \right|_2 = 2^n.
\end{align*}
\]

By (6.89), we obtain
\[
\begin{align*}
    x^n &= q(x)f(x) + b_{11}(x-1)(x-2) + b_{12}(x-2) + b_{21}(x-1)^2
\end{align*}
\]
and hence
\[
\begin{align*}
    A^n &= q(A)f(A) + b_{11}(A-I)(A-2I) + b_{12}(A-2I) + b_{21}(A-I)^2 \\
    &= b_{11}(A-I)(A-2I) + b_{12}(A-2I) + b_{21}(A-I)^2 \\
    &= b_{11} \begin{bmatrix} 2 & -2 & 5 \\ 1 & -1 & 0 \end{bmatrix} + b_{12} \begin{bmatrix} 1 & -2 & 5 \\ 1 & -2 & -1 \end{bmatrix} + b_{21} \begin{bmatrix} 2 & -2 & 5 \\ 1 & -1 & 0 \end{bmatrix} \\
    &= b_{11} \begin{bmatrix} * & * & * \\ 0 & 0 & 5 \end{bmatrix} + b_{12} \begin{bmatrix} * & * & * \\ 1 & -2 & 0 \end{bmatrix} + b_{21} \begin{bmatrix} * & * & * \\ 1 & -1 & 5 \end{bmatrix} \\
    &= \begin{bmatrix} * & * & * \\ b_{12} + b_{21} & -2b_{12} - b_{21} & 5(b_{11} + b_{21}) \end{bmatrix}.
\end{align*}
\]

Therefore,
\[
\begin{align*}
    \begin{bmatrix} a_{n+1} \\ a_n \\ 1 \end{bmatrix} &= A^n \begin{bmatrix} a_1 \\ a_0 \\ 1 \end{bmatrix} \\
    &= \begin{bmatrix} * & * & * \\ b_{12} + b_{21} & -2b_{12} - b_{21} & 5(b_{11} + b_{21}) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
    &= \begin{bmatrix} 5b_{11} - b_{12} + 5b_{21} \\ \end{bmatrix} \\
    \Rightarrow a_n &= 5b_{11} - b_{12} + 5b_{21} = -4 - 5n + 5(2^n).
\end{align*}
\]

2. Solutions of Linear ODEs with Constant Coefficients

As another application of JCF and minimal polynomial, let us consider the solutions of the following system of first order linear ordinary differential
equations (ODEs) with constant coefficients:

\[
\begin{align*}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + b_1 \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + b_2 \\
&\vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + b_n
\end{align*}
\]

(6.94)

where \(x_1, x_2, \ldots, x_n\) are unknown functions in \(t\) and \(a_{ij}\) and \(b_i\) are constants.

Using matrix notations, we may put it in the form

\[
\frac{d\begin{bmatrix} x_1 \\
x_2 \\
\vdots \\
x_n \end{bmatrix}}{dt} = \begin{bmatrix} A & b \end{bmatrix} \text{ or } \begin{bmatrix} x' \\
\end{bmatrix} = \begin{bmatrix} A & b \end{bmatrix}
\]

(6.95)

We call the system homogeneous if \(b = 0\) and non-homogeneous otherwise.

Note that we can always convert a non-homogeneous system of linear ODEs to a homogeneous one by introducing an extra unknown function \(x_{n+1}\) and turns (6.94) to

\[
\begin{align*}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + b_1x_{n+1} \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + b_2x_{n+1} \\
&\vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + b_nx_{n+1} \\
\frac{dx_{n+1}}{dt} &= 0
\end{align*}
\]

(6.96)

which is homogeneous. For a solution of (6.96) given by

\[
(x_1, x_2, \ldots, x_n, x_{n+1}) = (f_1(t), f_2(t), \ldots, f_n(t), f_{n+1}(t))
\]

(6.97)

\(f_{n+1}(t) \equiv c\) must be a constant. If \(f_{n+1}(t) \equiv 1\), then

\[
(x_1, x_2, \ldots, x_n) = (f_1(t), f_2(t), \ldots, f_n(t))
\]

(6.98)

is a solution of (6.94). More precisely,

\[
\{ \begin{bmatrix} x \\
x_{n+1} \end{bmatrix} : \frac{dx}{dt} = A \begin{bmatrix} x \\
x_{n+1} \end{bmatrix} + b \}
\]

(6.99)

\[
= \left\{ \begin{bmatrix} x \\
x_{n+1} \end{bmatrix} : \frac{d}{dt} \begin{bmatrix} x \\
x_{n+1} \end{bmatrix} = \begin{bmatrix} A & b \\
0 & 0 \end{bmatrix} \begin{bmatrix} x \\
x_{n+1} \end{bmatrix} \text{ and } x_{n+1} \equiv 1 \right\}.
\]
Therefore, to solve the non-homogeneous system (6.94), it suffices to solve (6.96) and choose its solutions with \(x_{n+1} \equiv 1\) as the solutions of the original system. So we only need to solve a homogeneous system \(x' = Ax\).

The solution of \(x' = Ax\) satisfying the initial condition \(x(t_0) = v\) is very explicitly given by

\[
(6.100) \quad x = e^{(t-t_0)A}v
\]

where \(e^B = \exp(B)\) is defined by

\[
(6.101) \quad e^B = \sum_{m=0}^{\infty} \frac{1}{m!}B^m
\]

for a square matrix \(B\). So it comes down to the computation of \(e^{(t-t_0)A}\), which can be achieved by the JCF or the minimal polynomial of \(A\).

It is easy to see that \(e^A\) has the following properties:

\[
(6.102) \begin{align*}
  e^{AI} &= e^{\lambda I} \\
  e^D &= \begin{bmatrix} e^{\lambda_1} \\ e^{\lambda_2} \\ \vdots \\ e^{\lambda_n} \end{bmatrix} \quad \text{for } D = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \\
  e^{-A} &= (e^A)^{-1} \\
  e^{A+B} &= e^A e^B \quad \text{if } AB = BA \\
  e^A v &= e^{\lambda} v \quad \text{if } Av = \lambda v \\
  P^{-1} e^A P &= \exp(P^{-1}AP).
\end{align*}
\]

Again, there are two ways to compute the solution \(e^{(t-t_0)A}v\) of \(x' = Ax\). The first approach uses the JCF of \(A\) and the second approach uses the minimal or characteristic polynomial of \(A\).

For simplicity, let us assume that \(t_0 = 0\). We may always replace \(t\) by \(t - t_0\) in \(e^{tA}v\) to obtain \(e^{(t-t_0)A}v\).

Suppose that \(P^{-1}AP = J\) is the JCF of \(A\) for an invertible matrix \(P\). Let \(B\) be the ordered basis of \(\mathbb{C}^n\) consisting of the column vectors of \(P\). As in (6.10), \(B = B_1 \sqcup B_2 \sqcup \ldots \sqcup B_l\) is a disjoint union of cycles of generalized eigenvectors. We write \(v\) as a linear combination of vectors in \(B\) as in (6.11):

\[
(6.103) \quad v = \sum_{i=1}^{l} \sum_{j=0}^{\infty} c_{ij}(A - \lambda_i I)^j v_i
\]

where \(v_1, v_2, \ldots, v_l\) are the vectors generating the cycles \(B_1, B_2, \ldots, B_l\) in \(B\).
We compute $e^{tA}v$ by writing $e^{tA} = e^{\lambda_i t}e^{(A-\lambda_i I)}$ for each $i$:

$$e^{tA}v = \sum_{i=1}^{l} e^{\lambda_i t} \sum_{j=0}^{\infty} c_{ij} (A-\lambda_i I)^j v_i$$

$$= \sum_{i=1}^{l} e^{\lambda_i t} e^{(A-\lambda_i I)} \sum_{j=0}^{\infty} c_{ij} (A-\lambda_i I)^j v_i$$

(6.104) $$= \sum_{i=1}^{l} e^{\lambda_i t} \left( \sum_{k=0}^{\infty} \frac{t^k}{k!} (A-\lambda_i I)^k \right) \left( \sum_{j=0}^{\infty} c_{ij} (A-\lambda_i I)^j v_i \right)$$

$$= \sum_{i=1}^{l} e^{\lambda_i t} \sum_{j,k \geq 0} \frac{c_{ij} t^k}{k!} (A-\lambda_i I)^{j+k} v_i.$$ 

Even that we sum from $j, k = 0$ to $\infty$, there are only finitely many terms among $(A-\lambda_i I)^{j+k} v_i$ that are nonzero since $(A-\lambda_i I)^m v_i = 0$ for $m_i = |B_i|$. In the special case that $A$ is diagonalizable, $m_i = 1$ for all $i$. That is, $(A-\lambda_i I)v_i = 0$ and then

$$e^{tA}v = \sum_{i=1}^{l} c_{i0} e^{\lambda_i t} v_i = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \ldots + c_l e^{\lambda_l t} v_l$$

(6.105) even that we sum from $j, k = 0$ to $\infty$, there are only finitely many terms among $(A-\lambda_i I)^{j+k} v_i$ that are nonzero since $(A-\lambda_i I)^m v_i = 0$ for $m_i = |B_i|$. In the special case that $A$ is diagonalizable, $m_i = 1$ for all $i$. That is, $(A-\lambda_i I)v_i = 0$ and then

$$e^{tA}v = \sum_{i=1}^{l} c_{i0} e^{\lambda_i t} v_i = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \ldots + c_l e^{\lambda_l t} v_l$$

(6.105) for $v = c_1 v_1 + c_2 v_2 + \ldots + c_l v_l$

where we let $c_i = c_{i0}$ and $v_1, v_2, \ldots, v_l$ are eigenvectors of $A$ that span $\mathbb{C}^n$.

**Example 6.8.** Let us solve the system of linear ODEs given by

$$\begin{cases}
\frac{dx_1}{dt} = 8x_1 - 4x_2 \\
\frac{dx_2}{dt} = 9x_1 - 4x_2
\end{cases} \quad \text{with} \quad \begin{cases}
x_1(0) = 1 \\
x_2(0) = 1
\end{cases}$$

(6.106) We put the system in the matrix form $\mathbf{x}' = A\mathbf{x}$ with

$$A = \begin{bmatrix} 8 & -4 \\ 9 & -4 \end{bmatrix}$$

(6.107) which is the same $2 \times 2$ matrix in Example 6.1. From the computation there, we know that

$$P^{-1}AP = J_{2,2} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{for} \quad P = [(A-2I)e_1 \ e_1] = \begin{bmatrix} 6 & 1 \\ 9 & 0 \end{bmatrix}.$$ 

(6.108) The solution of (6.106) is given by

$$\mathbf{x} = e^{tA} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = e^{tA} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$ 

(6.109)
We write $v$ as a linear combination of $e_1$ and $(A - 2I)e_1$:

$$v = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{9} \begin{bmatrix} 6 \\ 9 \end{bmatrix} = \frac{1}{3} e_1 + \frac{1}{9} (A - 2I)e_1. \quad (6.110)$$

Then

$$e^{tA}v = e^{tA} \left( \frac{1}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right)$$

$$= e^{2t}e^{t(A - 2I)} \left( \frac{1}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right)$$

$$= e^{2t} \left( \sum_{k=0}^{\infty} \frac{t^k}{k!} (A - 2I)^k \right) \left( \frac{1}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right)$$

$$= e^{2t} (1 + t(A - 2I)) \left( \frac{1}{3} e_1 + \frac{1}{9} (A - 2I)e_1 \right)$$

$$= e^{2t} \left( \frac{1}{3} e_1 + \frac{1}{9} \left(1 + \frac{t}{3}\right) (A - 2I)e_1 \right)$$

$$= e^{\frac{2t}{3}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + e^{\frac{2t}{3}} \left( \frac{1}{9} + \frac{t}{3} \right) \begin{bmatrix} 6 \\ 9 \end{bmatrix} = \left[ (1 + 2t) e^{2t} \right]. \quad (6.111)$$

So the solution is

$$\begin{cases} x_1 = (1 + 2t)e^{2t} \\ x_2 = (1 + 3t)e^{2t} \end{cases} \quad (6.112)$$

The difficulty with the above approach is the finding of the matrix $P$, just as in the computation of $A^n$. Our second approach will avoid this issue.

As before, we choose a nonzero polynomial $f(x)$ such that $f(A) = 0$ and try to divide $e^{tx}$ by $f(x)$. We may take $f(x)$ to be the characteristic or minimal polynomial of $A$.

Although $e^{tx}$ is not a polynomial, we can still divide $e^{tx}$ by $f(x)$ such that

$$e^{tx} = q(x)f(x) + r(x) \quad (6.113)$$

where $r(x)$ is a polynomial of $\deg r(x) < \deg f(x)$ and $q(x)$ is an entire function, i.e., a function that has complex derivatives everywhere for all $x \in \mathbb{C}$. When we replace $x$ by $A$, we still have

$$e^{tA} = q(A)f(A) + r(A). \quad (6.114)$$

So it suffices to find $r(x)$. As before, it can be computed using partial fractions:

$$\frac{e^{tx}}{f(x)} = q(x) + \frac{r(x)}{f(x)} = q(x) + \sum_{i=1}^{m} \sum_{j=1}^{d_i} \frac{c_{ij}}{(x - \lambda_i)^j} \quad (6.115)$$
for \( f(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \ldots (x - \lambda_m)^{d_m} \) with \( \lambda_1, \lambda_2, \ldots, \lambda_m \) distinct. As before, to determine \( c_{ij} \) for each \( i \), we multiply both sides of (6.115) by \((x - \lambda_i)^{d_i}\) and take derivatives at \( \lambda_i \) of order \( k = 0, 1, \ldots, d_i - 1 \), which yields

\[
(6.116) \quad c_{ij} = \left. \frac{(d_i - j)!}{f(x)} \frac{e^{tx}(x - \lambda_i)^{d_i}}{f(x)} \right|_{\lambda_i}.
\]

Example 6.9. Let us redo Example 6.8 using the characteristic polynomial of \( A \), which is \( f(x) = (x - 2)^2 \). Suppose that

\[
(6.117) \quad \frac{e^{tx}}{(x - 2)^2} = q(x) + \frac{c_1}{x - 2} + \frac{c_2}{(x - 2)^2}
\]

for some entire function \( q(x) \) and some constants \( c_1 \) and \( c_2 \). We multiply both sides by \((x - 2)^2\) and take derivatives of order \( k = 0, 1 \) at 2:

\[
(6.118) \quad c_1 = \left. (e^{tx})' \right|_{x=2} = te^{2t} \quad \text{and} \quad c_2 = \left. (e^{tx}) \right|_{x=2} = e^{2t}.
\]

Therefore,

\[
(6.119) \quad e^{tx} = q(x)(x - 2)^2 + c_1(x - 2) + c_2 = q(x)(x - 2)^2 + te^{2t}(x - 2) + e^{2t}.
\]

Replacing \( x \) by \( A \), we obtain

\[
(6.120) \quad e^{tA} = q(A)(A - 2I)^2 + te^{2t}(A - 2I) + e^{2t}I = te^{2t}(A - 2I) + e^{2t}.
\]

Then the solution of (6.106) is

\[
(6.121) \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = e^{tA}v = te^{2t}(A - 2I)v + e^{2t}v
\]

which agrees with (6.112).

Example 6.10. Let us solve the system of linear ODEs given by

\[
(6.122) \quad \begin{cases}
\frac{dx_1}{dt} = -2x_2 - x_3 \\
\frac{dx_2}{dt} = x_1 + 3x_2 + x_3 \quad \text{with} \quad \begin{cases} x_1(0) = 0 \\
\frac{dx_3}{dt} = x_1 + x_2 + 2x_3 \\
\end{cases}
\end{cases}
\]

We put it in the matrix form \( \mathbf{x}' = A\mathbf{x} \) for

\[
(6.123) \quad A = \begin{bmatrix} 0 & -2 & -1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}.
\]
The characteristic polynomial of $A$ is $f(x) = (x - 1)(x - 2)^2$. Suppose that

$$e^{tx}(x - 1)(x - 2)^2 = q(x) + \frac{a}{x - 1} + \frac{b_1}{x - 2} + \frac{b_2}{(x - 2)^2}$$

for some entire function $q(x)$ and some constants $a, b_1$ and $b_2$. Then

$$a = \left. \left( \frac{e^{tx}}{(x - 2)^2} \right) \right|_{x=1} = e^t,$$

$$b_1 = \left. \left( \frac{e^{tx}}{x - 1} \right) \right|_{x=2} = te^{2t} - e^{2t}, \text{ and}$$

$$b_2 = \left. \left( \frac{e^{tx}}{x - 1} \right) \right|_{x=2} = e^{2t}.$$