

12. LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

We now consider differential equations of the form

$$(12.1) \quad \dot{x} = Ax$$

with A an $n \times n$ real or complex matrix.

If $n = 1$, then we know the general solution has the form

$$x(t) = e^{At}x_0.$$

So, it is tempting to try to obtain a similar formula for the matrix case.

Consider the matrix power series

$$(12.2) \quad I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

The inequality $|AB| \leq |A| |B|$ for $n \times n$ matrices implies $|A^k| \leq |A|^k$. Hence the series (12.2) converges to a unique matrix. We call this matrix e^A or $\exp(A)$.

It is easy to see that if A and B commute (i.e., $AB = BA$), then $e^{A+B} = e^A \cdot e^B$.

Then, we can see that the matrix function e^{tA} defines a smooth function of t and, for any constant vector x_0 , the function

$$(12.3) \quad x(t) = e^{tA}x_0$$

solves the initial value problem $\dot{x} = Ax, x(0) = x_0$.

The matrix e^{tA} is a fundamental matrix for (12.1), since its columns consist of solutions (with initial values e_j , the standard unit vectors), and its determinant is nowhere zero.

The form (12.3) of the solution to (12.1) is useful for many purposes, but in some contexts it is useful to have other forms for the solutions to (12.1).

Suppose that the matrix A is diagonalizable, that is, there exists a nonsingular matrix X and n numbers $\lambda_1, \dots, \lambda_n$ such that

$$(12.4) \quad D = X^{-1}AX, \quad \text{or} \quad AX = XD,$$

where $D = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ is a diagonal matrix.

Denote $X = (x_1, \dots, x_n)$, where x_j is the j th column vector of X . The above equation gives that

$$A(x_1, \dots, x_n) = (x_1, \dots, x_n)D = (\lambda_1 x_1, \dots, \lambda_n x_n).$$

That is, the elements in the diagonal of D are the eigenvalues of A , and the column vectors of X are the corresponding eigenvectors of A .

The first equation in (12.4) also gives

$$e^{tD} = X^{-1}e^{tA}X, \quad \text{or} \quad e^{tA}X = Xe^{tD}.$$

Take the j th column of the last equality we have that

$$\phi_j(t) := e^{tA}x_j = e^{\lambda_j t}x_j$$

is a solutions of (12.1). Hence, $(\phi_1(t), \dots, \phi_n(t))$ form a form a fundamental set of solutions since the eigenvectors x_1, \dots, x_n are linearly independent. Thus, the general solution has the form

$$x(t) = \sum_j \alpha_j e^{t\lambda_j} x_j.$$

Now we suppose that the eigenvalues of A are all real, but that they are not necessarily distinct. Hence, A is not necessary diagonalizable.

Write the characteristic polynomial of A , $p(\lambda) = \det(\lambda I - A)$ as

$$p(\lambda) = (\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_s)^{m_s}$$

for some m_1, \dots, m_s . The real numbers $\lambda_1, \dots, \lambda_s$ are the distinct eigenvalues of A , and $\sum_j m_j = n$.

Denote

$$V_j = \ker(A - \lambda_j I)^{m_j}.$$

From linear algebra, we know that

$$p(A) = (A - \lambda_1 I)^{m_1} \dots (A - \lambda_s I)^{m_s} = 0.$$

That is, for any vector $v \in \mathbb{R}^n$, $p(A)v = 0$. Hence, there is a direct sum decomposition

$$\mathbb{R}^n = V_1 \oplus V_2 \oplus \dots \oplus V_s.$$

Also we have the following.

Fact 12.1. V_j have the following properties.

- (1) $A(V_j) \subseteq V_j$, that is, V_j is invariant under A .
- (2) $\dim V_j = m_j$.
- (3) For any $1 \leq j \leq s$, there is an integer $0 < r_j \leq m_j$ such that for each $v \in V_j$, $(A - \lambda_j I)^{r_j}(v) = 0$, and there is a $v \in V_j$ such that $(A - \lambda_j I)^{r_j-1}v \neq 0$.
- (4) V_j consists of only eigenvalues if and only if $r_j = 1$.

It is easy to see part (1), because if $v \in V_j$, then

$$(A - \lambda_j I)^{m_j} Av = A(A - \lambda_j I)^{m_j} v = A0 = 0,$$

hance $Av \in V_j$. Also, part (3) and (4) are obvious.

The subspace V_j is called the *generalized eigenspace* of A corresponding to the eigenvalue λ_j .

In the case that the matrix A is diagonalizable, each V_j is the eigenspace of λ_j and $r_j = 1$. In particular, if all the eigenvalues are distinct, then $\dim V_j = 1$ and $r_j = 1$ for any $1 \leq j \leq n$.

To find the general solution of (12.1), it suffices to find n linearly independent solutions. For this purpose, it suffices to work in each generalized eigenspace V_j .

For each j , denote $N_j = A - \lambda_j I$. Then V_j is the set of vectors v such that there is a positive integer $1 \leq r = r_j \leq m_j$ such that $N_j^r v = 0$. Note that $V_j \supsetneq \{0\}$ since the eigenvectors corresponding to λ_j are in V_j .

Lemma 12.2. *For any $v \in V_j$ with $N_j^{r_j-1}v \neq 0$, $v, N_j v, \dots, N_j^{r_j-1}v$ are linearly independent.*

Proof. Suppose

$$c_0 v + c_1 N_j v + \dots + c_{r_j-1} N_j^{r_j-1} v = 0.$$

Apply $N_j^{r_j-1}$ we get $c_0 N_j^{r_j-1} v = 0$. Since $N_j^{r_j-1} v \neq 0$, we have $c_0 = 0$. Hence we have

$$c_1 N_j v + \dots + c_{r_j-1} N_j^{r_j-1} v = 0.$$

Apply $N_j^{r_j-2}, \dots, N_j$, we get $c_1 = \dots = c_{r_j-2} = 0$, and then c_{r_j-1} . \square

Definition 12.1. *A subspace W of a vector space is called a cyclic subspace for a linear transformation N if there are a vector $w \in W$ and a positive integer r such that $N^{r-1}w \neq 0$, $N^r w = 0$, and $W = \text{span}(w, Nw, N^2w, \dots, N^{r-1}w)$.*

In this case, w is called a cyclic generator of W of order a .

We also call W a cyclic subspace of order r .

Remark. (1) *A cyclic subspace W is invariant under N . That is, $Nv \in W$ for any $v \in W$.*

(2) *If w is a cyclic generator of W of order r , then the vectors $w, Nw, \dots, N^{r-1}w$ form a basis for W . This is because by Lemma 12.2, the set of vectors form a linearly independent set.*

Definition 12.2. *An operator N is called nilpotent on a vector space V , if $N^a = 0$ for some integer $r > 0$. That is, there is a positive integer m such that $N^r v = 0$ for all $v \in V_j$.*

We will use the following theorem in algebra.

Theorem 12.3. *If N is a nilpotent operator on a finite dimensional vector space V , then V is a direct sum of cyclic subspaces. More precisely, there are cyclic subspaces W_1, W_2, \dots, W_k and positive integers r_1, \dots, r_k such that W_i is cyclic of order r_i and*

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_k.$$

Let us apply this theorem to the subspaces V_j and operators $N_j = A - \lambda_j I$ above. The operator N_j is nilpotent on V_j . So we may write

$$V_j = W_{j1} \oplus W_{j2} \oplus \dots \oplus W_{jk},$$

where each W_{ji} is cyclic of order r_i . Let w_i be a generator for N_j on W_{ji} . Then, the set $\{N_j^l w_i : 0 \leq l < r_i, 1 \leq i \leq k\}$ is a basis for V_j .

Next, let us indicate how we can obtain the general solution to (12.1) in a single cyclic subspace W_{ji} of N_j .

To simplify the notation, assume that $W = W_{ji}$, $m = r_i = \dim W$, $\lambda = \lambda_j$, $N = N_j|_W$. Then, $A|_W$ has only the eigenvalue λ , this eigenvalue has multiplicity m , and $N = A - \lambda I$. The vector $w = w_i$ is in W , and the vectors $w, Nw, N^2w, \dots, N^{m-1}w$ form a basis for W .

Lemma 12.4. *The functions*

$$x_{m-j}(t) = e^{t\lambda} \left(\sum_{k=m-j}^{m-1} \frac{t^{k-m+j}}{(k-m+j)!} N^k w \right)$$

for $j = 1, \dots, m$ form a fundamental set of solutions of $\dot{x} = Ax$ on W .

Before giving the proof, let us write the solutions

$$x_0(t), x_1(t), \dots, x_{m-1}(t)$$

in a more extended form:

$$\begin{aligned} x_{m-1}(t) &= e^{\lambda t} N^{m-1} w, \\ x_{m-2}(t) &= e^{\lambda t} (N^{m-2} w + t N^{m-1} w), \\ x_{m-3}(t) &= e^{\lambda t} \left(N^{m-3} w + t N^{m-2} w + \frac{t^2}{2!} N^{m-1} w \right), \\ &\vdots \\ x_0(t) &= e^{\lambda t} \left(w + t N w + \dots + \frac{t^{m-1}}{(m-1)!} N^{m-1} w \right). \end{aligned}$$

Proof of Lemma 12.4. The vectors $w, Nw, \dots, N^{m-1}w$ are linearly independent in W .

If we show the functions $x_{m-j}(t)$ are solutions, then it follows that they are linearly independent since their values at $t = 0$ are the independent vectors $w, Nw, \dots, N^{m-1}w$.

But, the function $x_{m-j}(t)$ has the form

$$x_{m-j}(t) = e^{t(\lambda I + N)} N^{m-j} w = e^{tA} N^{m-j} w,$$

which is a solution of (12.1). □

Remark 12.5. *In general to apply the above results one needs to first find a cyclic vector w of order r for the subspace W . If λ is an eigenvalue of multiplicity k of A , and $\ker(A - \lambda I)$ is one dimensional, then there is a simple procedure to find a cyclic vector for $A - \lambda I$ on W . Let w_0 be any eigenvector for A associated to λ which is in the image of $A - \lambda I$, and in general, if w_{i-1} is in $\ker(A - \lambda I)^i$ for some $i < k$, let w_i be such that $(A - \lambda I)w_i = w_{i-1}$. Then, w_{k-1} is a cyclic vector for $A - \lambda I$ on W , and $\{w_{k-1}, \dots, w_1, w_0\}$ form a basis for W .*

Now, suppose that A has some complex eigenvalues, but is a real matrix.

If μ is a complex eigenvalue of A , and the corresponding eigenvector is w , where w is a vector in \mathbb{C}^n , then the equation $\dot{z} = Az$ has complex solutions

$$z(t) = e^{\mu t}w \quad \text{and} \quad \bar{z}(t) = e^{\bar{\mu}t}\bar{w}.$$

Hence, the equation $\dot{x} = Ax$ has solutions

$$\operatorname{Re}(e^{\mu t}w) \quad \text{and} \quad \operatorname{Im}(e^{\mu t}w).$$

In general, if μ is a complex eigenvalue of multiplicity k , then we can decompose $W = \ker(A - \mu I)^k$ into cyclic subspaces and get basis $(w, Nw, \dots, N^{r-1}w)$ for each cyclic subspace, where r is the dimension of the cyclic subspace. Then we can get solutions $z_{r-1}(t), \dots, z_1(t), z_0(t)$. The real parts and imaginary parts of the solutions are the solutions to the equation $\dot{x} = Ax$.

Examples.

1.
$$\begin{aligned} x' &= -2x \\ y' &= y \end{aligned}$$

Here the matrix A is $\begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}$, the eigenvalues are $-2, 1$, and the general solution is

$$\bar{x}(t) = e^{-2t} \begin{pmatrix} c_1 \\ 0 \end{pmatrix} + e^t \begin{pmatrix} 0 \\ c_2 \end{pmatrix}.$$

The critical point 0 is called a *saddle*.

2.
$$\begin{aligned} x' &= 2x - y \\ y' &= x + y \end{aligned}$$

The matrix A is $\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$. The characteristic polynomial is $\lambda^2 - 3\lambda + 3$, and the eigenvalues are $\lambda = \frac{3}{2} \pm i\frac{\sqrt{3}}{2}$.

Letting $\lambda = \frac{3}{2} + i\frac{\sqrt{3}}{2}$, we have the matrix equation

$$(A - \lambda I) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This gives $(2 - \lambda)v_1 = v_2$, so that a complex eigenvalue is $(v_1, v_2) = (1, 2 - \lambda)$.

We get a complex solution of the form

$$\begin{aligned} \bar{x}_c(t) &= e^{\lambda t} \begin{pmatrix} 1 \\ 2 - \lambda \end{pmatrix} = e^{(\frac{3}{2} + i\frac{\sqrt{3}}{2}t)} \begin{pmatrix} 1 \\ \frac{1}{2} - i\frac{\sqrt{3}}{2} \end{pmatrix} \\ &= e^{(\frac{3}{2} + i\frac{\sqrt{3}}{2}t)} \left(\begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} + i \begin{bmatrix} 0 \\ -\frac{\sqrt{3}}{2} \end{bmatrix} \right). \end{aligned}$$

The real and imaginary parts of this are

$$\begin{aligned} \text{Re} &= e^{\frac{3}{2}t} \left(\cos\left(\frac{\sqrt{3}}{2}t\right) \begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} - \sin\left(\frac{\sqrt{3}}{2}t\right) \begin{bmatrix} 0 \\ -\frac{\sqrt{3}}{2} \end{bmatrix} \right), \\ \text{Im} &= e^{\frac{3}{2}t} \left(\cos\left(\frac{\sqrt{3}}{2}t\right) \begin{bmatrix} 0 \\ -\frac{\sqrt{3}}{2} \end{bmatrix} + \sin\left(\frac{\sqrt{3}}{2}t\right) \begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} \right). \end{aligned}$$

3.

$$\begin{aligned} x' &= 2x \\ y' &= 2y \end{aligned}$$

The matrix A is $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. The characteristic polynomial is $(\lambda - 2)^2$, and the only eigenvalue is 2.

The general solution is

$$x(t) = c_1 e^{2t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

4.

$$\begin{aligned} x' &= x + y \\ y' &= y \end{aligned}$$

The matrix is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We have

$$e^{tA} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} \left[I + t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right] = \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix}.$$

So the general solution is

$$\mathbf{x}(t) = e^{tA} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{bmatrix} c_1 e^t + c_2 t e^t \\ c_2 e^t \end{bmatrix}.$$

5.

$$x' = 3x + 11y + 5z$$

$$y' = -x - y - z$$

$$z' = 2x + z$$

The matrix is $A = \begin{pmatrix} 3 & 11 & 5 \\ -1 & -1 & -1 \\ 2 & 0 & 1 \end{pmatrix}$. The characteristic polynomial is

$$p(\lambda) = \lambda^3 - 3\lambda^2 + 4 = (\lambda - 2)^2(\lambda + 1).$$

The eigenvalue $\lambda = 2$:

Let $N = A - 2I$. Then, $N = \begin{bmatrix} 1 & 11 & 5 \\ -1 & -3 & -1 \\ 2 & 0 & -1 \end{bmatrix}$ with $\text{rank}(N) = 2$,

and $N^2 = \begin{bmatrix} 0 & -22 & -11 \\ 0 & -2 & -1 \\ 0 & 22 & 11 \end{bmatrix}$ with $\text{rank}(N)^2 = 1$.

The vector $v = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ is in $\ker(N)$. The vector $w = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ satisfies $Nw = v$. We get two linearly independent solutions in $\ker(N^2)$ by

$$e^{2t}v, \quad e^{2t}(w + tNw).$$

The eigenvalue $\lambda = -1$:

Let $N = A + I$. Then, $N = \begin{bmatrix} 4 & 11 & 5 \\ -1 & 0 & -1 \\ 2 & 0 & 2 \end{bmatrix}$, $\text{rank}(N) = 2$, and $\ker(N)$ is one-dimensional.

The vector $v = \begin{pmatrix} -1 \\ -\frac{1}{11} \\ 1 \end{pmatrix}$ is in the kernel of N , so is an eigenvector for A associated to $\lambda = -1$.

A fundamental set of solutions, then, is the set

$$e^{2t} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, e^{2t}(I + tN) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e^{-t} \begin{pmatrix} -1 \\ -\frac{1}{11} \\ 1 \end{pmatrix},$$

where

$$N = \begin{bmatrix} 1 & 11 & 5 \\ -1 & -3 & -1 \\ 2 & 0 & -1 \end{bmatrix}.$$

Appendix: Jordan canonical form

Recall that the characteristic polynomial of an $n \times n$ matrix A , $p(\lambda) = \det(\lambda I - A)$, can be written as

$$p(\lambda) = (\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_s)^{m_s},$$

for some m_1, \dots, m_s with $\sum_j m_j = n$, where the real numbers $\lambda_1, \dots, \lambda_s$ are the distinct eigenvalues of A .

If we denote $V_j = \ker(A - \lambda_j I)^{m_j}$, then there is a direct sum decomposition

$$\mathbb{R}^n = V_1 \oplus V_2 \oplus \dots \oplus V_s,$$

where $\dim V_j = m_j$. For each j , $1 \leq j \leq s$, V_j is a direct sum of cyclic subspaces of \mathbb{R}^n

$$V_j = W_{j1} \oplus W_{j2} \oplus \dots \oplus W_{jt_j},$$

where each W_{ji} is cyclic of order r_{ji} with respect to the operator $N_j = A - \lambda_j I$, and $\sum_i r_{ji} = m_j$.

Moreover, on each W_{ji} , there exists a vector $w = w_{ji}$, a cyclic generator, such that the set

$$(12.5) \quad \{N_j^{r_{ji}-1}w, \dots, N_j w, w\}$$

form a basis for W_{ji} . Hence, if we take this basis, the matrix representation of the operator is a Jordan block

$$(12.6) \quad B_{ji} = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_j & 1 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & \lambda_j & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_j \end{pmatrix}$$

restricted to W_{ji} . So if we take basis of the form (12.5) for each W_{ji} , then restricted to V_j , the matrix has the form

$$(12.7) \quad J_j = \begin{pmatrix} B_{j1} & 0 & \dots & 0 & 0 \\ 0 & B_{j2} & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & 0 & B_{jt_j} \end{pmatrix}.$$

A Jordan form of the matrix A is

$$J = \text{diag}(J_1, \dots, J_s).$$

If a matrix A has complex eigenvalues λ_j , then the Jordan form is a complex matrix. Now we consider the case that A is an $n \times n$ real matrix and it has complex eigenvalues λ_j .

The complex eigenvalues come in complex conjugate pairs

$$\mu_1, \bar{\mu}_1, \mu_2, \bar{\mu}_2, \dots, \mu_k, \bar{\mu}_k,$$

where $\mu_j = \lambda_j + i\kappa_j$, $\bar{\mu}_j = \lambda_j - i\kappa_j$, and $i = \sqrt{-1}$.

If $\mu = \lambda + i\kappa$ is a complex eigenvalue of A , with corresponding eigenvector $w = u + iv$, then $\bar{w} = u - iv$ is a eigenvector of A belongs to the eigenvalue $\mu = \lambda - i\kappa$. That is,

$$A(u + iv) = (\lambda + i\kappa)(u + iv), \quad A(u - iv) = (\lambda - i\kappa)(u - iv).$$

Hence we get that

$$\begin{aligned} Au &= \text{Re } A(u + iv) = \text{Re}(\lambda + i\kappa)(u + iv) = \lambda u - \kappa v, \\ Av &= \text{Im } A(u + iv) = \text{Im}(\lambda + i\kappa)(u + iv) = \kappa u + \lambda v. \end{aligned}$$

So if we take (u, v) as a basis in the subspace $W = \text{span}(v, u)$ spanned by u and v , then restricted to W , A has the form

$$(12.8) \quad \begin{pmatrix} \lambda & -\kappa \\ \kappa & \lambda \end{pmatrix}.$$

In general, if W is a cyclic subspace for $N = A - \mu I$, where $\mu = \lambda + i\kappa$ is a complex eigenvalue of A , then there is a cyclic subspace \bar{W} for $\bar{N} = A - \bar{\mu}I$ with $r = \dim N = \dim \bar{N}$. In this case, if $w = u + iv \in W$ is a cyclic generator of W for the nilpotent operator N , then $\bar{w} = u - iv \in \bar{W}$ is a cyclic generator of \bar{W} for \bar{N} . Hence we have $A(N^k w) = \mu N^k w + N^{k+1} w$ for $k = 1, \dots, r - 2$ and $A(N^{r-1} w) = \mu N^{r-1} w + N^r w = \mu N^{r-1} w$, and therefore

$$\begin{aligned} A(\text{Re } N^k w) &= \text{Re } A(N^k w) = \lambda \text{Re } N^k w - \kappa \text{Im } N^k w + \text{Re } N^{k+1} w, \\ A(\text{Im } N^k w) &= \text{Im } A(N^k w) = \kappa \text{Re } N^k w + \lambda \text{Im } N^k w + \text{Im } N^{k+1} w, \end{aligned}$$

for $k = 1, \dots, r-2$ and

$$A(\operatorname{Re} N^{r-1}w) = \operatorname{Re} A(N^{r-1}w) = \lambda \operatorname{Re} N^{r-1}w - \kappa \operatorname{Im} N^{r-1}w,$$

$$A(\operatorname{Im} N^{r-1}w) = \operatorname{Im} A(N^{r-1}w) = \kappa \operatorname{Re} N^{r-1}w + \lambda \operatorname{Im} N^{r-1}w.$$

So if we take

$$\{\operatorname{Re} N^{r-1}w, \operatorname{Im} N^{r-1}w, \dots, \operatorname{Re} Nw, \operatorname{Im} Nw, \operatorname{Re} w, \operatorname{Im} w\}$$

as a basis for $W \oplus \overline{W}$, then restricted to $W \oplus \overline{W}$, the Jordan block has the form

$$(12.9) \quad B_{ji} = \begin{pmatrix} M & I & 0 & \dots & 0 & 0 \\ 0 & M & I & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & M & I \\ 0 & 0 & 0 & \dots & 0 & M \end{pmatrix},$$

where $M = \begin{pmatrix} \lambda & -\kappa \\ \kappa & \lambda \end{pmatrix}$ and $I = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$.