Linear Differential Equations with Constant Coefficients

We now consider differential equations of the form

$$\dot{x} = Ax \tag{1}$$

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

If A were a scalar, 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

$$I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots \tag{2}$$

Because of the inequality $\mid AB \mid \leq \mid A \mid\mid B \mid$ for $n \times n$ matrices, the series (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

$$x(t) = e^{tA} x_0 \tag{3}$$

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

The matrix e^{tA} is a fundamental matrix for (1), since its columns consist of solutions (with initial values e_j, the standard unit vectors), and its determinant is nowhere zero.

The form (3) of the solution to (1) is useful for many purposes, but in some contexts it is useful to have other forms for the solutions to (1).

Suppose that the matrix A is real and diagonalizable over the reals with distinct real eigenvalues $\lambda_1, \dots, \lambda_n$. This means that there is a basis $\{v_1, \dots, v_n\}$ for $\mathbf{R}^n$ in which v_j is an eigenvector corresponding to λ_j.

Then, it is easy to see that the function $x_j(t) = e^{t\lambda_j} v_j$ is a solution, and these form a fundamental set of solutions. Thus, the general solution has the form

$$x(t) = \sum_j \alpha_j e^{t\lambda_j} v_j \tag{4}$$

Now suppose that the eigenvalues of A are all real, but that they are not necessarily distinct.

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

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

Thus, the real numbers $\lambda_1, \dots, \lambda_s$ are the distinct eigenvalues of A, and $\sum_j m_j = n$. From linear algebra, we know that there is a direct sum decomposition

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

where each $V_j = ker(A - \lambda_j I)^{mj}$.

Then, there is an integer $0 < r_j \leq m_j$

1. $A(V_j) \subseteq V_j$ (V_j is invariant)
2. The characteristic polynomial of A on V_j is $(\lambda - \lambda_j)^{m_j}$
3. For each $v \in V_j$, $(A - \lambda_j I)^{r_j}(v) = 0$
4. There is a $v \in V_j$ such that $(A - \lambda_j I)^{r_j - 1} v \neq 0$

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

In the case above in which the eigenvalues are distinct, each V_j is the eigenspace of λ_j and $r_j = 1$.

In general, each V_j contains an eigenvector. It consists only of eigenvectors (its non-zero elements of course) if and only if $r_j = 1$.

To find the general solution of (1), it suffices to find n linearly independent solutions. For this purpose, it suffices to work in each generalized eigenspace V_j. Indeed, each generalized eigenspace is invariant by the differential equation: any solution which starts in V_j remains in V_j for all t.

We remark that, once we have found the eigenvalues of A, there is a simple procedure to find the generalized eigenspaces. If λ_j is an eigenvalue, and we set $N_j = A - \lambda_j I$, then V_j is the set of vectors v such that there is a positive integer $m \geq 1$ such that $N_j^m v = 0$. Note that V_j contains non-zero vectors since the eigenvectors corresponding to λ_j are in V_j.

In fact, if we set $V_{jk} = ker(N_j^k)$, then we have the increasing family of subspaces

$$\{0\} \subsetneq V_{j1} \subseteq V_{j2} \subseteq \ldots \subseteq V_j$$

Since V_j is finite dimensional this increasing sequence of subspaces must actually be finite and we have an $r_j > 0$ such that

$$\{0\} \subsetneq V_{j1} \subseteq V_{j2} \subset \ldots \subseteq V_{j,r_j} = V_j.$$

The space V_{j1} is just the space spanned by the eigenvectors.

Let us now give a procedure which gets a basis for V_j.

The operator N_j is *nilpotent* on V_j. That is, there is a positive integer m such that $N_j^m v = 0$ for all $v \in V_j$.

There is a general structure theory for nilpotent operators. To recall this theory, assume that V is is a finite dimensional vector space and N is a nilpotent operator on V.

Given a finite set of vectors $w_1, w_2, \ldots w_k$ in V, let $sp(w_1, w_2, \ldots, w_k)$ denote the subspace spanned by the vectors w_i.

A subspace W of V is called a *cyclic subspace* if there are a vector $w \in W$ and a positve integer a such that $N^{a-1}w \neq 0, N^a w = 0$, and $W = sp(w, Nw, N^2 w, \ldots, N^{a-1}w)$. In that case, we call w a *cyclic generator* of W of *order* a. We also call W a cyclic subspace of order a.

Remarks.

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 a, then the vectors $w, Nw, \ldots, N^{a-1}w$ form a basis for W. To see this, it suffices to prove that these vectors form a linearly independent set.

   Suppose that we have constants $c_i, i = 0, \ldots, a-1$ such that

   $$v \stackrel{\text{def}}{=} c_0 w + c_1 N w + \ldots + c_{a-1} N^{a-1} w = 0.$$

   Then,

$$N^{a-1}v = 0 = c_0 N^{a-1}w$$

since all other terms involve $N^b w$ with $b \geq a$. Since, $N^{a-1}w \neq 0$, we have $c_0 = 0$.

Hence, the c_0 term is not actually in the expression for v. That is,

$$v = c_1 Nw + c_2 N^2 w + \ldots + c_{a-1} N^{a-1} w = 0.$$

Now, repeating the argument using $N^{a-2}v$, we get

$$c_1 N^{a-1} w = 0$$

so that $c_1 = 0$. Continuing with this argument, we get each $c_i = 0$ proving the required linear independence statement.

Theorem. *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 $W_1, W_2, \ldots, W_k$ and positive integers $a_1, \ldots, a_k$ such that W_i is cyclic of order a_i and*

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

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

$$V_j = W_{j1} \oplus W_{j2} \oplus \ldots \oplus W_{jk}$$

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

Next, let us indicate how we can obtain the general solution to (1) in a single cyclic subspace W_{ji} of N_j.

To simplify the notation, assume that $W = W_{ji},\ m = a_i = dim\ W, \lambda = \lambda_j,\ N = N_j \mid W$. Then, $A \mid W$ 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 $w, Nw, N^2 w, \ldots, N^{m-1} w$ form a basis for W.

Lemma. *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 $j = 1, \ldots, m-1$ *form a fundamental set of solutions of of* $\dot{x} = Ax$ *on* W.

Before giving the proof, let us write the solutions

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

in a more extended form.

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

Proof.
The vectors $w, Nw, \ldots, N^{m-1}w$ 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 $w, Nw, \ldots, N^{m-1}w$.

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 (1). QED.

Remark. *In general to apply the above results one needs to first find a cyclic vector* w *of order* a *for the subspace* W. *If* λ *is an eigenvalue of multiplicity two of* A, *and* $W = ker(A - \lambda I)^2$, *then there is a simple procedure to find a cyclic vector for* $A - \lambda I$ *on* W. *Let* v *be any eigenvector for* A

associated to λ which is in the image of $A - \lambda I$, and let w be such that $(A - \lambda I)w = v$. Then, w is a cyclic vector for $A - \lambda I$ on W.

Now, suppose that A has some complex eigenvalues, but is a real matrix. The complex eigenvalues come in complex conjugate pairs

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

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

We can consider the matrix as an operator on $\mathbf{C}^n$ by defining $A(u+iv) = Au + iAv$.

First suppose that A is complex diagonalizable with real eigenvalues $\lambda_1, \ldots, \lambda_j$ and complex eigenvalues

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

For each real eigenvalue λ_s, let $v_s^1, v_s^2, \ldots, v_s^{p_s}$ be a basis for $ker(A - \lambda_s I)$.

The functions $x_s^h(t) = e^{\lambda_s t} v_s^h$ form a set of linearly independent $\mathbf{R}^n$-valued solutions to (1).

For $s = 1, \ldots, k$, let $\zeta_s^1, \zeta_s^2, \ldots, \zeta_s^{q_s}$ be a complex eigenvector of A corresponding to μ_s. The functions $y_s^h(t) = e^{\mu_s t}\zeta_s^h$ form a set of linearly independent $\mathbf{C}^n$ valued solutions to the complexified differential equation $\dot{z} = Az$ on $\mathbf{C}^n$.

Then, the functions $x_s^h(t), Re(y_s^h(t)), Im(y_s^h(t))$ (note h and s vary) form a complete set of linearly independent solutions for (1) in $\mathbf{R}^n$.

In the general case, the real eigenvalues λ_s will have generalized eigenspaces (not just consisting of eigenvectors), and so will the complex eigenvalues μ_s. Moreover, the generalized eigenspace of $\bar{\mu}_s$ is the set of complex conjugates of elements of the generalized eigenspace of μ_s. One then finds a real basis $\{v_s^h\}$ for the generalized eigenspace of each λ_s and a complex basis $\{\zeta_s^h\}$ for the generalized eigenspace of each μ_s.

The solutions

$$x_s^h(t) = e^{\lambda_s t} \left(\sum_{\ell=0}^{p_s^h} \frac{t^\ell}{\ell!}(A - \lambda_s I)^\ell \right) v_s^h,$$

$$Re\left(e^{\mu_s t} \left(\sum_{\ell=0}^{q_s^h} \frac{t^\ell}{\ell!}(A - \mu_s I)^\ell \right) \zeta_s^h \right),$$

$$Im\left(e^{\mu_s t}\left(\sum_{\ell=0}^{q_s^h}\frac{t^\ell}{\ell!}(A-\mu_s I)^\ell\right)\zeta_s^h\right),$$

then form a set of n linearly independent $\mathbf{R}^n$ valued solutions to (1).

Note that, as a consequence of the above techniques, we get that all the components of solutions to (1) are linear combinations of terms of the form $t^k exp(t\lambda_j)$ where λ_j is a real eigenvalue or

$$t^k e^{a_j t} cos(b_j t),\ t^k e^{a_j t} sin(b_j t)$$

where $a_j + ib_j$ is a complex eigenvalue.