Section 9.8: The Matrix Exponential Function

Definition and Properties of Matrix Exponential

In the final section, we introduce a new notation which allows the formulas for solving normal systems with constant coefficients to be expressed identically to those for solving first-order equations with constant coefficients. For example, a general solution to $x'(t) = ax(t)$ where a is a constant is $x(t) = ce^{at}$. Similarly, a general solution to the normal system $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ where $\mathbf{A}$ is a constant $n \times n$ matrix is $\mathbf{x}(t) = \mathbf{c}e^{\mathbf{A}t}$. However, to see this we will need to define what the exponential notation means for a matrix. To do this, we generalize the Taylor series expansion of e^x.

Definition. If $\mathbf{A}$ is a constant $n \times n$ matrix, the matrix exponential $e^{\mathbf{A}t}$ is given by

$$(1) \qquad e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \cdots + \mathbf{A}^n\frac{t^n}{n!} + \cdots ,$$

where the right-hand side indicates the $n \times n$ matrix whose elements are power series with coefficients given by the entries in the matrices.

The exponential is easiest to compute when $\mathbf{A}$ is diagonal. For the matrix $\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$, we calculate

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \mathbf{A}^3 = \begin{bmatrix} -1 & 0 \\ 0 & 8 \end{bmatrix}, \ldots, \mathbf{A}^n = \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix}.$$

Then we get

$$e^{\mathbf{A}t} = \sum_{n=0}^{\infty} \mathbf{A}^n \frac{t^n}{n!} = \begin{bmatrix} \displaystyle\sum_{n=0}^{\infty}(-1)^n\frac{t^n}{n!} & 0 \\ 0 & \displaystyle\sum_{n=0}^{\infty} 2^n\frac{t^n}{n!} \end{bmatrix} = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{2t} \end{bmatrix}.$$

In general, if $\mathbf{A}$ is an $n \times n$ diagonal matrix with entries $r_1, r_2, \ldots, r_n$ on the main diagonal, then $e^{\mathbf{A}t}$ is the diagonal matrix with entries $e^{r_1 t}, e^{r_2 t}, \ldots, e^{r_n t}$ on the main diagonal. We will show later in the section how to calculate the matrix exponential for another class of matrices.

It turns out that the series (1) converges for all t and shares many properties with the scalar exponential e^{at}.

Theorem 1. *Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ constant matrices, and $r, s, t \in \mathbb{R}$. Then*

(1) $e^{\mathbf{A}0} = e^{\mathbf{0}} = \mathbf{I}$

(2) $e^{\mathbf{A}(t+s)} = e^{\mathbf{A}t}e^{\mathbf{A}s}$

(3) $(e^{\mathbf{A}t})^{-1} = e^{-\mathbf{A}t}$

(4) $e^{(\mathbf{A}+\mathbf{B})t} = e^{\mathbf{A}t}e^{\mathbf{B}t}$ *if* $\mathbf{AB} = \mathbf{BA}$

(5) $e^{r\mathbf{I}t} = e^{rt}\mathbf{I}$.

Matrix Exponential and Fundamental Matrices

Item (3) in Theorem 1 tells us that for any matrix $\mathbf{A}$, $e^{\mathbf{A}t}$ has an inverse for all t, and it is found by simply replacing t with $-t$. Another noteworthy property of the matrix exponential comes from differentiating the series (1) term by term:

$$\begin{aligned}\frac{d}{dt}(e^{\mathbf{A}t}) &= \frac{d}{dt}\left(\mathbf{I} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \cdots + \mathbf{A}^n\frac{t^n}{n!} + \cdots\right)\\ &= \mathbf{A} + \mathbf{A}^2 t + \mathbf{A}^3\frac{t^2}{2!} + \cdots + \mathbf{A}^n\frac{t^{n-1}}{(n-1)!} + \cdots\\ &= \mathbf{A}\left[\mathbf{I} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \cdots + \mathbf{A}^{n-1}\frac{t^{n-1}}{(n-1)!} + \cdots\right].\end{aligned}$$

Therefore, $\frac{d}{dt}(e^{\mathbf{A}t}) = \mathbf{A}e^{\mathbf{A}t}$, so $e^{\mathbf{A}t}$ is a solution to the matrix DE $\mathbf{X}' = \mathbf{AX}$. Moreover, the fact that $e^{\mathbf{A}t}$ is invertible implies that its columns are linearly independent solutions to the system $\mathbf{x}' = \mathbf{Ax}$. This leads to the following.

Theorem 2. *If* $\mathbf{A}$ *is an* $n \times n$ *constant matrix, then the columns of the matrix exponential* $e^{\mathbf{A}t}$ *form a fundamental solution set for the system* $\mathbf{x}'(t) = \mathbf{Ax}(t)$*. Therefore,* $e^{\mathbf{A}t}$ *is a fundamental matrix for the system, and a general solution is* $\mathbf{x}(t) = \mathbf{c}e^{\mathbf{A}t}$.

If we have already calculated a fundamental matrix for the system, this simplifies greatly the computation of the matrix exponential.

Theorem 3. *Let* $\mathbf{X}(t)$ *and* $\mathbf{Y}(t)$ *be two fundamental matrices for the same system* $\mathbf{x}' = \mathbf{Ax}$*. Then there exists a constant matrix* $\mathbf{C}$ *such that* $\mathbf{Y}(t) = \mathbf{X}(t)\mathbf{C}$ *for all* t*. In particular,*

$$e^{\mathbf{A}t} = \mathbf{X}(t)\mathbf{X}(0)^{-1}. \tag{2}$$

In the case that $\mathbf{A}$ has n linearly independent eigenvectors $\mathbf{u}_i$, then Theorem 3 tells us that

$$e^{\mathbf{A}t} = [e^{r_1 t}\mathbf{u}_1\ e^{r_2 t}\mathbf{u}_2\ \cdots\ e^{r_n t}\mathbf{u}_n][\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_n]^{-1}.$$

But what if the matrix does not have n linearly independent eigenvectors? So far, we do not have any way of finding a fundamental matrix for such a system. We now try to remedy this problem.

Nilpotent Matrices and Generalized Eigenvectors

Definition. A matrix $\mathbf{A}$ is nilpotent if $\mathbf{A}^k = \mathbf{0}$ for some positive integer k.

We introduce this class of matrices because the calculation of their exponential is simplified - it only has a finite number of terms:

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \cdots + \mathbf{A}^{k-1}\frac{t^{k-1}}{(k-1)!} + \mathbf{0} + \mathbf{0} + \cdots = \mathbf{I} + \mathbf{A}t + \cdots + \mathbf{A}^{k-1}\frac{t^{k-1}}{(k-1)!}.$$

Additionally, a theorem from linear algebra (Cayley-Hamilton Theorem) tells us that if $\mathbf{A}$ has only one eigenvalue r_1 (that is, the characteristic polynomial has the form $p(r) = (r - r_1)^n$), then $\mathbf{A} - r_1\mathbf{I}$ is nilpotent and $(\mathbf{A} - r_1\mathbf{I})^n = \mathbf{0}$, allowing us to write

$$e^{\mathbf{A}t} = e^{r_1\mathbf{I}t}e^{(\mathbf{A}-r_1\mathbf{I})t} = e^{r_1 t}\left[\mathbf{I} + (\mathbf{A} - r_1\mathbf{I})t + \cdots + (\mathbf{A} - r_1\mathbf{I})^{n-1}\frac{t^{n-1}}{(n-1)!}\right].$$

Example 1. Find the fundamental matrix $e^{\mathbf{A}t}$ for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where

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

Solution. We find the characteristic polynomial for $\mathbf{A}$:

$$p(r) = |\mathbf{A} - r\mathbf{I}| = \begin{vmatrix} 2-r & 1 & 1 \\ 1 & 2-r & 1 \\ -2 & -2 & -1-r \end{vmatrix} = -(r-1)^3.$$

Therefore, $r = 1$ is the only eigenvalue of $\mathbf{A}$, so $(\mathbf{A} - \mathbf{I})^3 = \mathbf{0}$ and

$$e^{\mathbf{A}t} = e^t e^{(\mathbf{A}-\mathbf{I})t} = e^t\left\{\mathbf{I} + (\mathbf{A} - \mathbf{I})t + (\mathbf{A} - \mathbf{I})^2\frac{t^2}{2}\right\}. \tag{3}$$

We calculate

$$\mathbf{A} - \mathbf{I} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & -2 \end{bmatrix} \text{ and } (\mathbf{A} - \mathbf{I})^2 = \mathbf{0}.$$

Substitution into (3) gives us

$$e^{\mathbf{A}t} = e^t\mathbf{I} + te^t(\mathbf{A} - \mathbf{I}) = \begin{bmatrix} e^t + te^t & te^t & te^t \\ te^t & e^t + te^t & te^t \\ -2te^t & -2te^t & e^t - 2te^t \end{bmatrix}. \qquad \diamondsuit$$

We would like to calculate fundamental matrices for any system with a constant coefficient matrix, not just when the matrix is nilpotent. The key to this is to generalize the concept of an eigenvector.

Definition. Let $\mathbf{A}$ be a square matrix. If a nonzero vector $\mathbf{u}$ satisfies the equation

$$(\mathbf{A} - r\mathbf{I})^m \mathbf{u} = \mathbf{0} \tag{4}$$

for some scalar r and positive integer m, we call $\mathbf{u}$ a generalized eigenvector associated with r.

Why are generalized eigenvectors useful? For one thing, they allow us to calculate $e^{\mathbf{A}t}\mathbf{u}$ with a finite number of terms *without* having to find $e^{\mathbf{A}t}$:

$$\begin{aligned} e^{\mathbf{A}t}\mathbf{u} &= e^{r\mathbf{I}t} e^{(\mathbf{A}-r\mathbf{I})t}\mathbf{u} \\ &= e^{rt}\left[\mathbf{I}\mathbf{u} + t(\mathbf{A}-r\mathbf{I})\mathbf{u} + \cdots + \frac{t^{m-1}}{(m-1)!}(\mathbf{A}-r\mathbf{I})^{m-1}\mathbf{u} + \frac{t^m}{m!}(\mathbf{A}-r\mathbf{I})^m\mathbf{u} + \cdots\right] \\ &= e^{rt}\left[\mathbf{u} + t(\mathbf{A}-r\mathbf{I})\mathbf{u} + \cdots + \frac{t^{m-1}}{(m-1)!}(\mathbf{A}-r\mathbf{I})^{m-1}\mathbf{u}\right]. \end{aligned}$$

Additionally, by Theorem 2 $e^{\mathbf{A}t}\mathbf{u}$ is a solution to the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$, so if we can find n generalized eigenvectors $\mathbf{u}_i$ for the $n \times n$ matrix $\mathbf{A}$ which are linearly independent, the corresponding solutions $\mathbf{x}_i(t) = e^{\mathbf{A}t}\mathbf{u}_i$ will be a fundamental solution set and thus we'll have a fundamental matrix. Then by Theorem 3 we can write the matrix exponential:

$$e^{\mathbf{A}t} = \mathbf{X}(t)\mathbf{X}(0)^{-1} = [e^{\mathbf{A}t}\mathbf{u}_1 \ e^{\mathbf{A}t}\mathbf{u}_2 \ \cdots \ e^{\mathbf{A}t}\mathbf{u}_n][\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n]^{-1}. \tag{5}$$

It is not hard to see that since any regular eigenvector is also a generalized eigenvector, if $\mathbf{A}$ has a full set of n linearly independent eigenvectors, then the above representation (5) is exactly the one we get from the methods of previous sections. Returning to our earlier question, what about when $\mathbf{A}$ is defective - when it has fewer than n linearly independent eigenvectors? The following theorem from linear algebra tells us that the new method works even in this case.

Theorem 4. *(Primary Decomposition Theorem) Let* $\mathbf{A}$ *be a constant* $n \times n$ *matrix with characteristic polynomial*

$$p(r) = (r - r_1)^{m_1}(r - r_2)^{m_2} \cdots (r - r_k)^{m_k},$$

where the r_i*'s are the distinct eigenvalues of* $\mathbf{A}$. *Then for each* i *there exist* m_i *linearly independent generalized eigenvectors satisfying*

$$(\mathbf{A} - r_i\mathbf{I})^{m_i}\mathbf{u} = \mathbf{0}.$$

Moreover, $m_1 + m_2 + \cdots + m_k = n$ *and the full collection of these* n *generalized eigenvectors is linearly independent.*

We summarize below the procedure for finding a fundamental solution set for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for any constant square matrix $\mathbf{A}$.

(1) Calculate the characteristic polynomial $p(r) = |\mathbf{A} - r\mathbf{I}|$ and express it in the form $p(r) = (r-r_1)^{m_1}(r-r_2)^{m_2}\cdots(r-r_k)^{m_k}$, where $r_1, r_2, \ldots, r_k$ are the distinct zeros.

(2) For each eigenvalue r_i, find m_i linearly independent generalized eigenvectors by solving the system $(\mathbf{A} - r_i\mathbf{I})^{m_i}\mathbf{u} = \mathbf{0}$.

(3) Form n linearly independent solutions by finding

$$\mathbf{x}(t) = e^{rt}\left[\mathbf{u} + t(\mathbf{A} - r\mathbf{I})\mathbf{u} + \frac{t^2}{2!}(\mathbf{A} - r\mathbf{I})^2\mathbf{u} + \cdots\right]$$

for each generalized eigenvector $\mathbf{u}$ found in part (2); for the eigenvalue r_i, the series terminates after at most m_i terms.

(4) Assemble the fundamental matrix $\mathbf{X}(t)$ from the n solutions and obtain the exponential $e^{\mathbf{A}t}$ using (5).

Example 2. Find the fundamental matrix $e^{\mathbf{A}t}$ for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where

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

Solution. Start by finding the characteristic polynomial:

$$p(r) = |\mathbf{A} - r\mathbf{I}| = \begin{vmatrix} 1-r & 0 & 0 \\ 1 & 3-r & 0 \\ 0 & 1 & 1-r \end{vmatrix} = -(r-1)^2(r-3).$$

So the eigenvalues are $r = 1$ with multiplicity 2 and $r = 3$ with multiplicity 1.
For $r = 1$, we find 2 linearly independent generalized eigenvectors satisfying $(\mathbf{A}-\mathbf{I})^2\mathbf{u} = \mathbf{0}$:

$$(\mathbf{A} - \mathbf{I})^2\mathbf{u} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 4 & 0 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

implies that $u_1 = -2u_2$, so by setting $u_2 = s, u_3 = v$ we obtain the eigenvectors

$$\mathbf{u} = \begin{bmatrix} -2s \\ s \\ v \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + v\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For each of these, we calculate a solution vector:

$$\begin{aligned} \mathbf{x}_1(t) &= e^t[\mathbf{u}_1 + t(\mathbf{A} - \mathbf{I})\mathbf{u}_1] \\ &= e^t\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + te^t\begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \\ &= e^t\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + te^t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2e^t \\ e^t \\ te^t \end{bmatrix}. \end{aligned}$$

$$\mathbf{x}_2(t) = e^t[\mathbf{u}_2 + t(\mathbf{A} - \mathbf{I})\mathbf{u}_2] = e^t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + te^t\begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ e^t \end{bmatrix}.$$

For $r = 3$, we find a generalized eigenvector satisfying $(\mathbf{A} - 3\mathbf{I})\mathbf{u} = \mathbf{0}$:

$$\begin{bmatrix} -2 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

implies $u_1 = 0, u_2 = 2u_3$, so we have the eigenvector $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ and a third solution vector

$$\mathbf{x}_3(t) = e^{3t}\mathbf{u}_3 = e^{3t} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2e^{3t} \\ e^{3t} \end{bmatrix}.$$

Combining these solution vectors gives the fundamental matrix $\mathbf{X}(t) = \begin{bmatrix} 0 & -2e^t & 0 \\ 0 & e^t & 2e^{3t} \\ e^t & te^t & e^{3t} \end{bmatrix}$.

We set $t = 0$ and compute the inverse matrix:

$$\mathbf{X}(0) = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{X}(0)^{-1} = \begin{bmatrix} -\frac{1}{4} & -\frac{1}{2} & 1 \\ -\frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 \end{bmatrix}.$$

Finally, by equation (2) we have

$$\begin{aligned} e^{\mathbf{A}t} &= \mathbf{X}(t)\mathbf{X}(0)^{-1} = \begin{bmatrix} 0 & -2e^t & 0 \\ 0 & e^t & 2e^{3t} \\ e^t & te^t & e^{3t} \end{bmatrix} \begin{bmatrix} -\frac{1}{4} & -\frac{1}{2} & 1 \\ -\frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 \end{bmatrix} \\ &= \begin{bmatrix} e^t & 0 & 0 \\ -\frac{1}{2}e^t + \frac{1}{2}e^{3t} & e^{3t} & 0 \\ -\frac{1}{4}e^t - \frac{1}{2}te^t + \frac{1}{4}e^{3t} & -\frac{1}{2}e^t + \frac{1}{2}e^{3t} & e^t \end{bmatrix}. \quad \diamond \end{aligned}$$

As a closing remark, we note that use of the matrix exponential as a fundamental matrix simplifies many computations. For one example, the variation of parameters formula in Section 9.7 can be rewritten as

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}_0 + \int_{t_0}^{t} e^{\mathbf{A}(t-s)}\mathbf{f}(s)ds,$$

which more closely resembles the formula for the scalar case.

Homework: p. 557, #1-11 odd, 17-21 odd.