LS.6 Solution Matrices

In the literature, solutions to linear systems often are expressed using square matrices rather than vectors. You need to get used to the terminology. As before, we state the definitions and results for a 2×2 system, but they generalize immediately to $n \times n$ systems.

1. Fundamental matrices. We return to the system

(1) $$\mathbf{x}' = A(t)\,\mathbf{x}\ ,$$

with the general solution

(2) $$\mathbf{x} = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t)\ ,$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are two independent solutions to (1), and c_1 and c_2 are arbitrary constants.

We form the matrix whose columns are the solutions $\mathbf{x}_1$ and $\mathbf{x}_2$:

(3) $$X(t) = (\,\mathbf{x}_1 \quad \mathbf{x}_2\,) = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} .$$

Since the solutions are linearly independent, we called them in LS.5 a *fundamental* set of solutions, and therefore we call the matrix in (3) a **fundamental matrix** for the system (1).

Writing the general solution using X(t). As a first application of $X(t)$, we can use it to write the general solution (2) efficiently. For according to (2), it is

$$\mathbf{x} = c_1\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + c_2\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} ,$$

which becomes using the fundamental matrix

(4) $$\mathbf{x} = X(t)\,\mathbf{c} \qquad \text{where } \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} , \qquad (\textit{general solution to (1)}).$$

Note that the vector $\mathbf{c}$ must be written on the right, even though the c's are usually written on the left when they are the coefficients of the solutions $\mathbf{x}_i$.

Solving the IVP using X(t). We can now write down the solution to the IVP

(5) $$\mathbf{x}' = A(t)\,\mathbf{x}\ , \qquad \mathbf{x}(t_0) = \mathbf{x}_0.$$

Starting from the general solution (4), we have to choose the $\mathbf{c}$ so that the initial condition in (6) is satisfied. Substituting t_0 into (5) gives us the matrix equation for $\mathbf{c}$:

$$X(t_0)\,\mathbf{c} = \mathbf{x}_0\ .$$

Since the determinant $|X(t_0)|$ is the value at t_0 of the Wronskian of $\mathbf{x}_1$ amd $\mathbf{x}_2$, it is non-zero since the two solutions are linearly independent (Theorem 5.2C). Therefore the inverse matrix exists (by LS.1), and the matrix equation above can be solved for $\mathbf{c}$:

$$\mathbf{c} = X(t_0)^{-1}\mathbf{x}_0\ ;$$

using the above value of $\mathbf{c}$ in (4), the solution to the IVP (1) can now be written

(6) $$\mathbf{x} = X(t)X(t_0)^{-1}\mathbf{x}_0\ .$$

Note that when the solution is written in this form, it's "obvious" that $\mathbf{x}(t_0) = \mathbf{x}_0$, i.e., that the initial condition in (5) is satisfied.

An equation for fundamental matrices We have been saying "a" rather than "the" fundamental matrix since the system (1) doesn't have a unique fundamental matrix: thare are many different ways to pick two independent solutions of $\mathbf{x}' = A\,\mathbf{x}$ to form the columns of X. It is therefore useful to have a way of recognizing a fundamental matrix when you see one. The following theorem is good for this; we'll need it shortly.

Theorem 6.1 *$X(t)$ is a fundamental matrix for the system (1) if its determinant $|X(t)|$ is non-zero and it satisfies the matrix equation*

$$X' = A\,X\,, \tag{7}$$

where X' means that each entry of X has been differentiated.

Proof. Since $|X| \not\equiv 0$, its columns $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent, by section LS.5. And writing $X = (\,\mathbf{x}_1 \quad \mathbf{x}_2\,)$, (7) becomes, according to the rules for matrix multiplication,

$$(\,\mathbf{x}_1' \quad \mathbf{x}_2'\,) = A\,(\,\mathbf{x}_1 \quad \mathbf{x}_2\,) = (\,A\mathbf{x}_1 \quad A\mathbf{x}_2\,)\,,$$

which shows that

$$\mathbf{x}_1' = A\,x_1 \qquad \text{and} \qquad \mathbf{x}_2' = A\,\mathbf{x}_2\ ;$$

this last line says that $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions to the system (1). □

2. The normalized fundamental matrix.

Is there a "best" choice for fundamental matrix?

There are two common choices, each with its advantages. If the ODE system has constant coefficients, and its eigenvalues are real and distinct, then a natural choice for the fundamental matrix would be the one whose columns are the normal modes — the solutions of the form

$$\mathbf{x}_i = \vec{\alpha}_i e^{\lambda_i t}\,, \qquad i = 1, 2.$$

There is another choice however which is suggested by (6) and which is particularly useful in showing how the solution depends on the initial conditions. Suppose we pick $X(t)$ so that

$$X(t_0) = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,. \tag{8}$$

Referring to the definition (3), this means the solutions $\mathbf{x}_1$ and $\mathbf{x}_2$ are picked so

$$\mathbf{x}_1(t_0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\,, \qquad \mathbf{x}_2(t_0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\,. \tag{8$'$}$$

Since the $\mathbf{x}_i(t)$ are uniquely determined by these initial conditions, the fundamental matrix $X(t)$ satisfying (8) is also unique; we give it a name.

Definition 6.2 The unique matrix $\widetilde{X}_{t_0}(t)$ satisfying

$$\widetilde{X}_{t_0}' = A\,\widetilde{X}_{t_0}\,, \qquad \widetilde{X}_{t_0}(t_0) = I \tag{9}$$

is called the **normalized fundamental matrix** at t_0 for A.

For convenience in use, the definition uses Theorem 6.1 to guarantee $\widetilde{X}_{t_0}$ will actually be a fundamental matrix; the condition $|\widetilde{X}_{t_0}(t)| \neq 0$ in Theorem 6.1 is satisfied, since the definition implies $|\widetilde{X}_{t_0}(t_0)| = 1$.

To keep the notation simple, we will assume in the rest of this section that $t_0 = 0$, as it almost always is; then $\widetilde{X}_0$ is the normalized fundamental matrix. Since $\widetilde{X}_0(0) = I$, we get from (6) the matrix form for the solution to an IVP:

(10) *The solution to the IVP* $\quad \mathbf{x}' = A(t)\,\mathbf{x}, \;\; \mathbf{x}(0) = \mathbf{x}_0 \quad$ *is* $\quad \mathbf{x}(t) = \widetilde{X}_0(t)\mathbf{x}_0.$

Calculating $\widetilde{X}_0$. One way is to find the two solutions in (8′), and use them as the columns of $\widetilde{X}_0$. This is fine if the two solutions can be determined by inspection.

If not, a simpler method is this: find any fundamental matrix $X(t)$; then

$$\widetilde{X}_0(t) \;=\; X(t)\,X(0)^{-1}. \tag{11}$$

To verify this, we have to see that the matrix on the right of (11) satisfies the two conditions in Definition 6.2. The second is trivial; the first is easy using the rule for matrix differentiation:

If $M = M(t)$ and $B,\ C$ are constant matrices, then $\;(BM)' = BM',\;\; (MC)' = M'C,$

from which we see that since X is a fundamental matrix,

$$(X(t)X(0)^{-1})' = X(t)'X(0)^{-1} = AX(t)X(0)^{-1} = A(X(t)X(0)^{-1}),$$

showing that $X(t)X(0)^{-1}$ also satisfies the first condition in Definition 6.2. □

Example 6.2A Find the solution to the IVP: $\mathbf{x}' \;=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{x}\,, \quad \mathbf{x}(0) = \mathbf{x}_0\,.$

Solution Since the system is $x' = y,\;\; y' = -x$, we can find by inspection the fundamental set of solutions satisfying (8′) :

$$\begin{matrix} x = \;\;\cos t \\ y = -\sin t \end{matrix} \qquad \text{and} \qquad \begin{matrix} x = \sin t \\ y = \cos t \end{matrix}\;.$$

Thus by (10) the normalized fundamental matrix at 0 and solution to the IVP is

$$\mathbf{x} \;=\; \widetilde{X}\,\mathbf{x}_0 \;=\; \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \;=\; x_0 \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} + y_0 \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}\;.$$

Example 6.2B Give the normalized fundamental matrix at 0 for $\mathbf{x}' \;=\; \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix} \mathbf{x}\,.$

Solution. This time the solutions (8′) cannot be obtained by inspection, so we use the second method. We calculated the normal modes for this sytem at the beginning of LS.2; using them as the columns of a fundamental matrix gives us

$$X(t) \;=\; \begin{pmatrix} 3e^{2t} & -e^{-2t} \\ e^{2t} & e^{-2t} \end{pmatrix}\;.$$

Using (11) and the formula for calculating the inverse matrix given in LS.1, we get

$$X(0) = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}, \qquad X(0)^{-1} = \tfrac{1}{4}\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix},$$

so that

$$\widetilde{X}(t) = \tfrac{1}{4}\begin{pmatrix} 3e^{2t} & -e^{-2t} \\ e^{2t} & e^{-2t} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} = \tfrac{1}{4}\begin{pmatrix} 3e^{2t}+e^{2t} & 3e^{2t}-3e^{-2t} \\ e^{2t}-e^{-2t} & e^{2t}+3e^{-2t} \end{pmatrix}.$$

6.3 The Exponential matrix.

The work in the preceding section with fundamental matrices was valid for any linear homogeneous square system of ODE's,

$$\mathbf{x}' = A(t)\,\mathbf{x}.$$

However, if the system has *constant coefficients*, i.e., the matrix A is a constant matrix, the results are usually expressed by using the exponential matrix, which we now define.

Recall that if x is any real number, then

$$(12) \qquad e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots .$$

Definition 6.3 Given an $n \times n$ constant matrix A, the **exponential matrix** e^A is the $n \times n$ matrix defined by

$$(13) \qquad e^A = I + A + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!} + \ldots .$$

Each term on the right side of (13) is an $n \times n$ matrix; adding up the ij-th entry of each of these matrices gives you an infinite series whose sum is the ij-th entry of e^A. (The series always converges.)

In the applications, an independent variable t is usually included:

$$(14) \qquad e^{At} = I + At + A^2\,\frac{t^2}{2!} + \ldots + A^n\,\frac{t^n}{n!} + \ldots .$$

This is not a new definition, it's just (13) above applied to the matrix At in which every element of A has been multiplied by t, since for example

$$(At)^2 = At \cdot At = A \cdot A \cdot t^2 = A^2t^2.$$

Try out (13) and (14) on these two examples; the first is worked out in your book (Example 2, p. 417); the second is easy, since it is not an infinite series.

Example 6.3A Let $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, show: $e^A = \begin{pmatrix} e^a & 0 \\ 0 & e^b \end{pmatrix}$; $e^{At} = \begin{pmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{pmatrix}$

Example 6.3B Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, show: $e^A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$; $e^{At} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$

What's the point of the exponential matrix? The answer is given by the theorem below, which says that the exponential matrix provides a royal road to the solution of a square system with constant coefficients: no eigenvectors, no eigenvalues, you just write down the answer!

Theorem 6.3 *Let* A *be a square constant matrix. Then*

(15) *(a)* $e^{At} = \widetilde{X}_0(t)$, *the normalized fundamental matrix at 0;*

(16) *(b) the unique solution to the IVP* $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = \mathbf{x}_0$ *is* $\mathbf{x} = e^{At}\mathbf{x}_0$.

Proof. Statement (16) follows immediately from (15), in view of (10).

We prove (15) is true, by using the description of a normalized fundamental matrix given in Definition 6.2: letting $X = e^{At}$, we must show $X' = AX$ and $X(0) = I$.

The second of these follows from substituting $t = 0$ into the infinite series definition (14) for e^{At}.

To show $X' = AX$, we assume that we can differentiate the series (14) term-by-term; then we have for the individual terms

$$\frac{d}{dt} A^n \frac{t^n}{n!} = A^n \cdot \frac{t^{n-1}}{(n-1)!} ,$$

since A^n is a constant matrix. Differentiating (14) term-by-term then gives

$$(18) \qquad \begin{aligned} \frac{dX}{dt} = \frac{d}{dt} e^{At} &= A + A^2 t + \ldots + A^n \frac{t^{n-1}}{(n-1)!} + \ldots \\ &= A\, e^{At} = A\, X . \end{aligned}$$

Calculation of e^{At}.

The main use of the exponential matrix is in (16) — writing down explicitly the solution to an IVP. If e^{At} has to be actually calculated for a specific system, several techniques are available.

a) In simple cases, it can be calculated directly as an infinite series of matrices.

b) It can always be calculated, according to Theorem 6.3, as the normalized fundamental matrix $\widetilde{X}_0(t)$, using (11): $\widetilde{X}_0(t) = X(t)X(0)^{-1}$.

c) A third technique uses the exponential law

$$(19) \qquad e^{(B+C)t} = e^{Bt}e^{Ct}, \qquad \textit{valid if } BC = CB.$$

To use it, one looks for constant matrices B and C such that

$$(20) \qquad A = B + C, \qquad BC = CB, \qquad e^{Bt} \text{ and } e^{Ct} \text{ are computable;}$$

then

$$(21) \qquad e^{At} = e^{B\,t}e^{C\,t}.$$

Example 6.3C Let $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Solve $\mathbf{x}' = A\,\mathbf{x}$, $\mathbf{x}(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, using e^{At}.

Solution. We set $B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; then (20) is satisfied, and

$$e^{At} = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{2t} \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

by (21) and Examples 6.3A and 6.3B. Therefore, by (16), we get

$$\mathbf{x} = e^{At}\,\mathbf{x}_0 = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = e^{2t} \begin{pmatrix} 1+2t \\ 2 \end{pmatrix} .$$

Exercises: Sections 4G,H