PEARSON NEW INTERNATIONAL EDITION

Algebra

Michael Artin

Second Edition

Pearson New International Edition

Algebra

Michael Artin
Second Edition

PEARSON

Many elementary properties of differentiation carry over to matrix-valued functions. The product rule, whose proof is an exercise, is an example:

Lemma 5.3.6 Product Rule.

(a) Let $A(t)$ and $B(t)$ be differentiable matrix-valued functions of t, of suitable sizes so that their product is defined. Then the matrix product $A(t)B(t)$ is differentiable, and its derivative is

$$\frac{d(AB)}{dt} = \frac{dA}{dt}B + A\frac{dB}{dt}.$$

(b) Let $A_1, \ldots, A_k$ be differentiable matrix-valued functions of t, of suitable sizes so that their product is defined. Then the matrix product $A_1 \cdots A_k$ is differentiable, and its derivative is

$$\frac{d}{dt}(A_1 \cdots A_k) = \sum_{i=1}^{k} A_1 \cdots A_{i-1}\left(\frac{dA_i}{dt}\right)A_{i+1} \cdots A_k. \qquad \square$$

A system of homogeneous linear, first-order, constant-coefficient differential equations is a matrix equation of the form

(5.3.7) $$\frac{dX}{dt} = AX,$$

where A is a constant $n \times n$ matrix and $X(t)$ is an n-dimensional vector-valued function. Writing out such a system, we obtain a system of n differential equations

(5.3.8)
$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}\,x_1(t) \quad + \quad \cdots \quad + \quad a_{1n}\,x_n(t)\\
&\;\;\vdots\\
\frac{dx_n}{dt} &= a_{n1}\,x_1(t) \quad + \quad \cdots \quad + \quad a_{nn}\,x_n(t).
\end{aligned}$$

The $x_i(t)$ are unknown functions, and the scalars a_{ij} are given. For example, if

(5.3.9) $$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix},$$

(5.3.7) becomes a system of two equations in two unknowns:

(5.3.10)
$$\begin{aligned}
\frac{dx_1}{dt} &= 3x_1 + 2x_2\\
\frac{dx_2}{dt} &= x_1 + 4x_2\,.
\end{aligned}$$

The simplest systems are those in which A is a diagonal matrix with diagonal entries λ_i. Then equation (5.3.8) reads

(5.3.11) $$\frac{dx_i}{dt} = \lambda_i x_i(t), \quad i = 1, \ldots, n.$$

Here the unknown functions x_i are not mixed up by the equations, so we can solve for each one separately:

(5.3.12) $$x_i = c_i e^{\lambda_i t},$$

for some arbitrary constants c_i.

The observation that allows us to solve the differential equation (5.3.7) in many cases is this: If V is an eigenvector for A with eigenvalue λ, i.e., if $AV = \lambda V$, then

(5.3.13) $$X = e^{\lambda t} V$$

is a particular solution of (5.3.7). Here $e^{\lambda t}V$ must be interpreted as the product of the variable scalar $e^{\lambda t}$ and the constant vector V. Differentiation operates on the scalar function, fixing V, while multiplication by A operates on the vector V, fixing the scalar $e^{\lambda t}$. Thus $\frac{d}{dt}e^{\lambda t}V = \lambda e^{\lambda t}V$ and also $Ae^{\lambda t}V = \lambda e^{\lambda t}V$. For example,

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

are eigenvectors of the matrix (5.3.9), with eigenvalue 5 and 2, respectively, and

(5.3.14) $$\begin{bmatrix} e^{5t} \\ e^{5t} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2e^{2t} \\ -e^{2t} \end{bmatrix}$$

solve the system (5.3.10).

This observation allows us to solve (5.3.7) whenever the matrix A has distinct real eigenvalues. In that case every solution will be a linear combination of the special solutions (5.3.13). To work this out, it is convenient to diagonalize.

Proposition 5.3.15 Let A be an $n \times n$ matrix, and let P be an invertible matrix such that $\Lambda = P^{-1}AP$ is diagonal, with diagonal entries $\lambda_1, \ldots, \lambda_n$. The general solution of the system $\frac{dX}{dt} = AX$ is $X = P\tilde{X}$, where $\tilde{X} = (c_1 e^{\lambda_1 t}, \ldots, c_n e^{\lambda_n t})^t$ solves the equation $\frac{d\tilde{X}}{dt} = \Lambda\tilde{X}$.

The coefficients c_i are arbitrary. They are often determined by assigning *initial conditions* – the value of X at some particular t_0.

Proof. We multiply the equation $\frac{d\tilde{X}}{dt} = \Lambda\tilde{X}$ by P: $P\frac{d\tilde{X}}{dt} = P\Lambda\tilde{X} = AP\tilde{X}$. But since P is constant, $P\frac{d\tilde{X}}{dt} = \frac{d(P\tilde{X})}{dt} = \frac{dX}{dt}$. Thus $\frac{dX}{dt} = AX$. This reasoning can be reversed, so $\tilde{X}$ solves the equation with Λ if and only if X solves the equation with A. □

The matrix that diagonalizes the matrix (5.3.10) was computed before (4.6.8):

(5.3.16) $$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \quad \text{and} \quad \Lambda = \begin{bmatrix} 5 & \\ & 2 \end{bmatrix}.$$

Thus

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P\tilde{X} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 e^{5t} \\ c_2 e^{2t} \end{bmatrix} = \begin{bmatrix} c_1 e^{5t} + 2c_2 e^{2t} \\ c_1 e^{5t} - c_2 e^{2t} \end{bmatrix}. \tag{5.3.17}$$

In other words, every solution is a linear combination of the two basic solutions (5.3.14).

We now consider the case that the coefficient matrix A has distinct eigenvalues, but that they are not all real. To copy the method used above, we first consider differential equations of the form (5.3.1), in which a is a complex number. Properly interpreted, the solutions of such a differential equation still have the form ce^{at}. The only thing to remember is that e^{at} will now be a complex-valued function of the real variable t.

The definition of the derivative of a complex-valued function is the same as for real-valued functions, provided that the limit (5.3.5) exists. There are no new features. We can write any such function $x(t)$ in terms of its real and imaginary parts, which will be real-valued functions, say

$$x(t) = p(t) + iq(t). \tag{5.3.18}$$

Then x is differentiable if and only if p and q are differentiable, and if they are, the derivative of x is $p' + iq'$. This follows directly from the definition. The usual rules for differentiation, such as the product rule, hold for complex-valued functions. These rules can be proved either by applying the corresponding theorem for real functions to p and q, or by copying the proof for real functions.

The exponential of a complex number $a = r + si$ is defined to be

$$e^a = e^{r+si} = e^r(\cos s + i \sin s). \tag{5.3.19}$$

Differentiation of this formula shows that $de^{at}/dt = ae^{at}$. Therefore ce^{at} solves the differential equation (5.3.1), and the proof given at the beginning of the section shows that these are the only solutions.

Having extended the case of one equation to complex coefficients, we can use diagonalization to solve a system of equations (5.3.7) when A is a complex matrix with distinct eigenvalues.

For example, let $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. The vectors $v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ and $v_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$ are eigenvectors, with eigenvalues $1 + i$ and $1 - i$, respectively. Let $\mathbf{B}$ denote the basis (v_1, v_2). Then A is diagonalized by the matrix $P = [\mathbf{B}]$:

$$P^{-1}AP = \frac{1}{2}\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = \begin{bmatrix} 1+i & \\ & 1-i \end{bmatrix} = \Lambda. \tag{5.3.20}$$

Then $\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} c_1 e^{(1+i)t} \\ c_2 e^{(1-i)t} \end{bmatrix}$. The solutions of (5.3.7) are

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P\tilde{X} = \begin{bmatrix} c_1 e^{(1+i)t} + i c_2 e^{(1-i)t} \\ i c_1 e^{(1+i)t} + c_2 e^{(1-i)t} \end{bmatrix}, \tag{5.3.21}$$

where c_1, c_2 are arbitrary complex numbers. So every solution is a linear combination of the two basic solutions

$$\begin{bmatrix} e^{(1+i)t} \\ i e^{(1+i)t} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} i e^{(1-i)t} \\ e^{(1-i)t} \end{bmatrix}. \tag{5.3.22}$$

However, these solutions aren't very satisfactory, because we began with a system of differential equations with real coefficients, and the answer we obtained is complex. When the equation is real, we will want the real solutions. We note the following lemma:

Lemma 5.3.23 Let A be a real $n \times n$ matrix, and let $X(t)$ be a complex-valued solution of the differential equation $\frac{dX}{dt} = AX$. The real and imaginary parts of $X(t)$ solve the same equation. □

Now *every* solution of the original equation (5.3.7), whether real or complex, has the form (5.3.21) for some complex numbers c_i. So the real solutions are among those we have found. To write them down explicitly, we may take the real and imaginary parts of the complex solutions.

The real and imaginary parts of the basic solutions (5.3.22) are determined using (5.3.19). They are

$$\begin{bmatrix} e^t \cos t \\ -e^t \sin t \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} e^t \sin t \\ e^t \cos t \end{bmatrix}. \tag{5.3.24}$$

Every real solution is a real linear combination of these particular solutions.

5.4 THE MATRIX EXPONENTIAL

Systems of first-order linear, constant-coefficient differential equations can be solved formally, using the *matrix exponential*.

The exponential of an $n \times n$ real or complex matrix A is the matrix obtained by substituting A for x and I for 1 into the Taylor's series for e^x, which is

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots , \tag{5.4.1}$$

Thus by definition,

$$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots . \tag{5.4.2}$$

We will be interested mainly in the matrix valued function e^{tA} of the variable scalar t, so we substitute tA for A:

$$e^{tA} = I + \frac{tA}{1!} + \frac{t^2A^2}{2!} + \frac{t^3A^3}{3!} + \cdots . \tag{5.4.3}$$

Theorem 5.4.4

(a) The series (5.4.2) converges absolutely and uniformly on bounded sets of complex matrices.

(b) e^{tA} is a differentiable function of t, and its derivative is the matrix product Ae^{tA}.

(c) Let A and B be complex $n\times n$ matrices that commute: $AB = BA$. Then $e^{A+B} = e^A e^B$.

In order not to break up the discussion, we have moved the proof of this theorem to the end of the section.

The hypothesis that A and B commute is essential for carrying the fundamental property $e^{x+y} = e^x e^y$ over to matrices. Nevertheless, **(c)** is very useful.

Corollary 5.4.5 For any $n\times n$ complex matrix A, the exponential e^A is invertible, and its inverse is e^{-A}.

Proof. Because A and $-A$ commute, $e^A e^{-A} = e^{A-A} = e^0 = I$. □

Since matrix multiplication is relatively complicated, it is often not easy to write down the entries of the matrix e^A. They won't be obtained by exponentiating the entries of A unless A is a diagonal matrix. If A is diagonal, with diagonal entries $\lambda_1, \ldots, \lambda_n$, then inspection of the series shows that e^A is also diagonal, and that its diagonal entries are e^{λ_i}.

The exponential is also fairly easy to compute for a triangular 2×2 matrix. For example, if

$$A = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix},$$

then

$$e^A = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \frac{1}{1!}\begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} 1 & 3 \\ & 4 \end{bmatrix} + \cdots = \begin{bmatrix} e & * \\ & e^2 \end{bmatrix}. \tag{5.4.6}$$

It is a good exercise to calculate the missing entry $*$ directly from the series.

The exponential of e^A can be determined whenever we know a matrix P such that $\Lambda = P^{-1}AP$ is diagonal. Using the rule $P^{-1}A^kP = (P^{-1}AP)^k$ (4.6.12) and the distributive law for matrix multiplication,

$$P^{-1}e^AP = (P^{-1}IP) + \frac{(P^{-1}AP)}{1!} + \frac{(P^{-1}AP)^2}{2!} + \cdots = e^{P^{-1}AP} = e^{\Lambda}. \tag{5.4.7}$$

Suppose that Λ is diagonal, with diagonal entries λ_i. Then e^{Λ} is also diagonal, and its diagonal entries are e^{λ_i}. In this case we can compute e^A explicitly:

$$e^A = Pe^{\Lambda}P^{-1}. \tag{5.4.8}$$

For example, if $A = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$, then $P^{-1}AP = \Lambda = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$. So

$$e^A = Pe^{\Lambda}P^{-1} = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}\begin{bmatrix} e & \\ & e^2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix} = \begin{bmatrix} e & e^2-e \\ & e^2 \end{bmatrix}.$$

The next theorem relates the matrix exponential to differential equations:

Theorem 5.4.9 Let A be a real or complex $n\times n$ matrix. The columns of the matrix e^{tA} form a basis for the space of solutions of the differential equation $\frac{dX}{dt} = AX$.

Proof. Theorem 5.4.4**(b)** shows that the columns of e^{tA} solve the differential equation. To show that every solution is a linear combination of the columns, we copy the proof given at the beginning of Section 5.3. Let $X(t)$ be an arbitrary solution. We differentiate the matrix product $e^{-tA}X(t)$ using the product rule (5.3.6):

$$\frac{d}{dt}\left(e^{-tA}X(t)\right) = \left(-Ae^{-tA}\right)X(t) + e^{-tA}\left(AX(t)\right). \tag{5.4.10}$$

Fortunately, A and e^{-tA} commute. This follows directly from the definition of the exponential. So the derivative is zero. Therefore $e^{-tA}X(t)$ is a constant column vector, say $C = (c_1, \ldots, c_n)^t$, and $X(t) = e^{tA}C$. This expresses $X(t)$ as a linear combination of the columns of e^{tA}, with coefficients c_i. The expression is unique because e^{tA} is an invertible matrix. □

Though the matrix exponential always solves the differential equation (5.3.7), it may not be easy to apply in a concrete situation because computation of the exponential can be difficult. But if A is diagonalizable, the exponential can be computed as in (5.4.8). We can use this method of evaluating e^{tA} to solve equation (5.3.7). Of course we will get the same solutions as we did before. Thus if A, P, and Λ are as in (5.3.16), then

$$e^{tA} = Pe^{t\Lambda}P^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{5t} & \\ & e^{2t} \end{bmatrix}\left(-\frac{1}{3}\right)\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} (e^{5t}+2e^{2t}) & (2e^{5t}-2e^{2t}) \\ (e^{5t}-e^{2t}) & (2e^{5t}+e^{2t}) \end{bmatrix}.$$