

Matrix exponential

From Wikipedia, the free encyclopedia

In mathematics, the **matrix exponential** is a matrix function on square matrices analogous to the ordinary exponential function. Abstractly, the matrix exponential gives the connection between a matrix Lie algebra and the corresponding Lie group.

Let X be an $n \times n$ real or complex matrix. The exponential of X , denoted by e^X or $\exp(X)$, is the $n \times n$ matrix given by the power series

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k$$

where X^0 is defined to be the identity matrix I with the same dimensions as X .^[1]

The above series always converges, so the exponential of X is well-defined. If X is a 1×1 matrix the matrix exponential of X is a 1×1 matrix whose single element is the ordinary exponential of the single element of X .

Contents

- 1 Properties
 - 1.1 Linear differential equation systems
 - 1.2 The exponential of sums
 - 1.2.1 Golden–Thompson inequality
 - 1.2.2 Lieb's theorem
 - 1.3 The exponential map
 - 1.4 The determinant of the matrix exponential
- 2 Computing the matrix exponential
 - 2.1 Diagonalizable case
 - 2.2 Projection case
 - 2.3 Rotation case
 - 2.4 Nilpotent case
 - 2.5 Generalization
 - 2.6 Evaluation by Laurent series
 - 2.7 Evaluation by implementation of Sylvester's formula
- 3 Illustrations
- 4 Applications
 - 4.1 Linear differential equations
 - 4.1.1 Example (homogeneous)
 - 4.1.2 Example (inhomogeneous)
 - 4.2 Inhomogeneous case generalization: variation of parameters
- 5 Matrix-matrix exponentials
- 6 See also
- 7 References
- 8 External links

Properties

Let X and Y be $n \times n$ complex matrices and let a and b be arbitrary complex numbers. We denote the $n \times n$ identity matrix by I and the zero matrix by 0 . The matrix exponential satisfies the following properties:^[2]

- $e^0 = I$
- $e^{aX}e^{bX} = e^{(a+b)X}$
- $e^Xe^{-X} = I$
- If $XY = YX$ then $e^Xe^Y = e^Ye^X = e^{(X+Y)}$.
- If Y is invertible then $e^{YXY^{-1}} = Ye^XY^{-1}$.
- $\exp(X^T) = (\exp X)^T$, where X^T denotes the transpose of X . It follows that if X is symmetric then e^X is also symmetric, and that if X is skew-symmetric then e^X is orthogonal.
- $\exp(X^*) = (\exp X)^*$, where X^* denotes the conjugate transpose of X . It follows that if X is Hermitian then e^X is also Hermitian, and that if X is skew-Hermitian then e^X is unitary.
- A Laplace transform of matrix exponentials amounts to the resolvent, $\int_0^\infty dt e^{-ts} e^{tX} = I / (sI - X)$.

Linear differential equation systems

One of the reasons for the importance of the matrix exponential is that it can be used to solve systems of linear ordinary differential equations. The solution of

$$\frac{d}{dt}y(t) = Ay(t), \quad y(0) = y_0,$$

where A is a constant matrix, is given by

$$y(t) = e^{At}y_0.$$

The matrix exponential can also be used to solve the inhomogeneous equation

$$\frac{d}{dt}y(t) = Ay(t) + z(t), \quad y(0) = y_0.$$

See the section on applications below for examples.

There is no closed-form solution for differential equations of the form

$$\frac{d}{dt}y(t) = A(t)y(t), \quad y(0) = y_0,$$

where A is not constant, but the Magnus series gives the solution as an infinite sum.

The exponential of sums

For any real numbers (scalars) x and y we know that the exponential function satisfies $e^{x+y} = e^x e^y$. The same is true for commuting matrices. If matrices X and Y commute (meaning that $XY = YX$), then

$$e^{X+Y} = e^X e^Y.$$

However, for matrices that do not commute the above equality does not necessarily hold. In this case the Baker–Campbell–Hausdorff formula can be used to calculate e^{X+Y} .

The converse is not true in general. The equation $e^{X+Y} = e^X e^Y$ does not imply that X and Y commute.

For Hermitian matrices there are two notable theorems related to the trace of matrix exponentials.

Golden–Thompson inequality

If A and H are Hermitian matrices, then

$$\mathrm{tr} \exp(A + H) \leq \mathrm{tr}(\exp(A) \exp(H)).^{[3]}$$

Note that there is no requirement of commutativity. There are counterexamples to show that the Golden–Thompson inequality cannot be extended to three matrices – and, in any event, $\mathrm{tr}(\exp(A)\exp(B)\exp(C))$ is not guaranteed to be real for Hermitian A, B, C . However, the next theorem accomplishes this in one sense.

Lieb's theorem

The **Lieb's theorem**, named after Elliott H. Lieb, states that, for a fixed Hermitian matrix H , the function

$$f(A) = \mathrm{tr} \exp(H + \log A)$$

is concave on the cone of positive-definite matrices.^[4]

The exponential map

Note that the exponential of a matrix is always an invertible matrix. The inverse matrix of e^X is given by e^{-X} . This is analogous to the fact that the exponential of a complex number is always nonzero. The matrix exponential then gives us a map

$$\exp: M_n(\mathbb{C}) \rightarrow \mathrm{GL}(n, \mathbb{C})$$

from the space of all $n \times n$ matrices to the general linear group of degree n , i.e. the group of all $n \times n$ invertible matrices. In fact, this map is surjective which means that every invertible matrix can be written as the exponential of some other matrix^[5] (for this, it is essential to consider the field \mathbb{C} of complex numbers and not \mathbb{R}).

For any two matrices X and Y ,

$$\|e^{X+Y} - e^X\| \leq \|Y\| e^{\|X\|} e^{\|Y\|},$$

where $\|\cdot\|$ denotes an arbitrary matrix norm. It follows that the exponential map is continuous and Lipschitz continuous on compact subsets of $M_n(\mathbb{C})$.

The map

$$t \mapsto e^{tX}, \quad t \in \mathbb{R}$$

defines a smooth curve in the general linear group which passes through the identity element at $t = 0$.

In fact, this gives a one-parameter subgroup of the general linear group since

$$e^{tX} e^{sX} = e^{(t+s)X}.$$

The derivative of this curve (or tangent vector) at a point t is given by

$$\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X. \quad (1)$$

The derivative at $t = 0$ is just the matrix X , which is to say that X generates this one-parameter subgroup.

More generally,^[6] for a generic t -dependent exponent, $X(t)$,

$$\frac{d}{dt}e^{X(t)} = \int_0^1 e^{\alpha X(t)} \frac{dX(t)}{dt} e^{(1-\alpha)X(t)} d\alpha.$$

Taking the above expression $e^{X(t)}$ outside the integral sign and expanding the integrand with the help of the Hadamard lemma one can obtain the following useful expression for the derivative of the matrix exponent,

$$\left(\frac{d}{dt}e^{X(t)}\right)e^{-X(t)} = \frac{d}{dt}X(t) + \frac{1}{2!}[X(t), \frac{d}{dt}X(t)] + \frac{1}{3!}[X(t), [X(t), \frac{d}{dt}X(t)]] + \dots$$

Note that the coefficients in the expression above are different from what appears in the exponential. For a closed form, see derivative of the exponential map.

The determinant of the matrix exponential

By Jacobi's formula, for any complex square matrix the following trace identity holds:^[7]

$$\det(e^A) = e^{\text{tr}(A)}.$$

In addition to providing a computational tool, this formula demonstrates that a matrix exponential is always an invertible matrix. This follows from the fact that the right hand side of the above equation is always non-zero, and so $\det(e^A) \neq 0$, which implies that e^A must be invertible.

In the real-valued case, the formula also exhibits the map

$$\exp: M_n(\mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R})$$

to not be surjective, in contrast to the complex case mentioned earlier. This follows from the fact that, for real-valued matrices, the right-hand side of the formula is always positive, while there exist invertible matrices with a negative determinant.

Computing the matrix exponential

Finding reliable and accurate methods to compute the matrix exponential is difficult, and this is still a topic of considerable current research in mathematics and numerical analysis. Matlab, GNU Octave, and SciPy all use the Padé approximant.^{[8][9][10]}

Diagonalizable case

If a matrix is diagonal:

$$A = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{bmatrix},$$

then its exponential can be obtained by exponentiating each entry on the main diagonal:

$$e^A = \begin{bmatrix} e^{a_1} & 0 & \dots & 0 \\ 0 & e^{a_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{a_n} \end{bmatrix}.$$

This also allows one to exponentiate diagonalizable matrices. If $A = UDU^{-1}$ and D is diagonal, then $e^A = Ue^DU^{-1}$. Application of Sylvester's formula yields the same result. (To see this, note that addition and multiplication, hence also exponentiation, of diagonal matrices is equivalent to element-wise addition and multiplication, and hence exponentiation; in particular, the "one-dimensional" exponentiation is felt element-wise for the diagonal case.)

Projection case

If P is a projection matrix (i.e. is idempotent: $P^2=P$, so multiplying P by itself any number of times is itself), its matrix exponential is $e^P = I + (e - 1)P$. This may be derived by expansion of the definition of the exponential function and by use of the idempotency of P :

$$e^P = \sum_{k=0}^{\infty} \frac{P^k}{k!} = I + \left(\sum_{k=1}^{\infty} \frac{1}{k!} \right) P = I + (e - 1)P.$$

Rotation case

For a simple rotation in which the perpendicular unit vectors a and b specify a plane,^[11] the rotation matrix R can be expressed in terms of a similar exponential function involving a generator G and angle θ .^{[12][13]}

$$\begin{aligned} G &= ba^T - ab^T & a^T b &= 0 \\ -G^2 &= aa^T + bb^T = P & P^2 &= P & PG &= GP = G, \end{aligned}$$

$$\begin{aligned} R(\theta) &= e^{G\theta} = I + G \sin(\theta) + G^2(1 - \cos(\theta)) \\ &= I - P + P \cos(\theta) + G \sin(\theta). \end{aligned}$$

The formula for the exponential results from reducing the powers of G in the series expansion and identifying the respective series coefficients of G^2 and G with $-\cos(\theta)$ and $\sin(\theta)$ respectively. The second expression here for $e^{G\theta}$ is the same as the expression for $R(\theta)$ in the article containing the derivation of the generator, $R(\theta) = e^{G\theta}$.

In two dimensions, if $a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, then $G = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $G^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, and

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = I \cos(\theta) + G \sin(\theta)$$

reduces to the standard matrix for a plane rotation.

The matrix $P = -G^2$ projects a vector onto the ab -plane and the rotation only affects this part of the vector. An example illustrating this is a rotation of $30^\circ = \pi/6$ in the plane spanned by a and b ,

$$a = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad b = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad G = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

$$P = -G^2 = \frac{1}{5} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix} \quad P \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 5 \\ 8 \\ 16 \end{pmatrix} = a + \frac{8}{\sqrt{5}}b$$

$$\theta = \frac{\pi}{6} \quad \Rightarrow \quad R = \frac{1}{10} \begin{pmatrix} 5\sqrt{3} & -\sqrt{5} & -2\sqrt{5} \\ \sqrt{5} & 8 + \sqrt{3} & -4 + 2\sqrt{3} \\ 2\sqrt{5} & -4 + 2\sqrt{3} & 2 + 4\sqrt{3} \end{pmatrix}$$

Let $N = I - P$, so $N^2 = N$ and its products with P and G are zero. This will allow us to evaluate powers of R .

$$R\left(\frac{\pi}{6}\right) = N + P\frac{\sqrt{3}}{2} + G\frac{1}{2} \quad R\left(\frac{\pi}{6}\right)^2 = N + P\frac{1}{2} + G\frac{\sqrt{3}}{2}$$

$$R\left(\frac{\pi}{6}\right)^3 = N + G \quad R\left(\frac{\pi}{6}\right)^6 = N - P \quad R\left(\frac{\pi}{6}\right)^{12} = N + P = I$$

Nilpotent case

A matrix N is nilpotent if $N^q = 0$ for some integer q . In this case, the matrix exponential e^N can be computed directly from the series expansion, as the series terminates after a finite number of terms:

$$e^N = I + N + \frac{1}{2}N^2 + \frac{1}{6}N^3 + \cdots + \frac{1}{(q-1)!}N^{q-1}.$$

Generalization

When the minimal polynomial of a matrix X can be factored into a product of first degree polynomials, it can be expressed as a sum

$$X = A + N$$

where

- A is diagonalizable
- N is nilpotent
- A commutes with N (i.e. $AN = NA$)

This is the Jordan–Chevalley decomposition.

This means that we can compute the exponential of X by reducing to the previous two cases:

$$e^X = e^{A+N} = e^A e^N.$$

Note that we need the commutativity of A and N for the last step to work.

Another (closely related) method if the field is algebraically closed is to work with the Jordan form of X . Suppose that $X = PJP^{-1}$ where J is the Jordan form of X . Then

$$e^X = Pe^J P^{-1}.$$

Also, since

$$\begin{aligned} J &= J_{a_1}(\lambda_1) \oplus J_{a_2}(\lambda_2) \oplus \cdots \oplus J_{a_n}(\lambda_n), \\ e^J &= \exp(J_{a_1}(\lambda_1) \oplus J_{a_2}(\lambda_2) \oplus \cdots \oplus J_{a_n}(\lambda_n)) \\ &= \exp(J_{a_1}(\lambda_1)) \oplus \exp(J_{a_2}(\lambda_2)) \oplus \cdots \oplus \exp(J_{a_k}(\lambda_k)). \end{aligned}$$

Therefore, we need only know how to compute the matrix exponential of a Jordan block. But each Jordan block is of the form

$$J_a(\lambda) = \lambda I + N$$

where N is a special nilpotent matrix. The matrix exponential of this block is given by

$$e^{\lambda I + N} = e^\lambda e^N.$$

Evaluation by Laurent series

By virtue of the Cayley–Hamilton theorem the matrix exponential is expressible as a polynomial of order $n-1$.

If P and Q_t are nonzero polynomials in one variable, such that $P(A) = 0$, and if the meromorphic function

$$f(z) = \frac{e^{tz} - Q_t(z)}{P(z)}$$

is entire, then

$$e^{tA} = Q_t(A).$$

To prove this, multiply the first of the two above equalities by $P(z)$ and replace z by A .

Such a polynomial $Q_t(z)$ can be found as follows—see Sylvester's formula. Letting a be a root of P , $Q_{a,t}(z)$ is solved from the product of P by the principal part of the Laurent series of f at a : It is proportional to the relevant Frobenius covariant. Then the sum S_t of the $Q_{a,t}$, where a runs over all the roots of P , can be taken as a particular Q_t . All the other Q_t will be obtained by adding a multiple of P to $S_t(z)$. In particular, $S_t(z)$, the Lagrange–Sylvester polynomial, is the only Q_t whose degree is less than that of P .

Example: Consider the case of an arbitrary 2-by-2 matrix,

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The exponential matrix e^{tA} , by virtue of the Cayley–Hamilton theorem, must be of the form

$$e^{tA} = s_0(t) I + s_1(t) A.$$

(For any complex number z and any \mathbf{C} -algebra B , we denote again by z the product of z by the unit of B .)

Let α and β be the roots of the characteristic polynomial of A ,

$$P(z) = z^2 - (a + d) z + ad - bc = (z - \alpha)(z - \beta) .$$

Then we have

$$S_t(z) = e^{\alpha t} \frac{z - \beta}{\alpha - \beta} + e^{\beta t} \frac{z - \alpha}{\beta - \alpha} ,$$

hence

$$s_0(t) = \frac{\alpha e^{\beta t} - \beta e^{\alpha t}}{\alpha - \beta}, \quad s_1(t) = \frac{e^{\alpha t} - e^{\beta t}}{\alpha - \beta}$$

if $\alpha \neq \beta$; while, if $\alpha = \beta$,

$$S_t(z) = e^{\alpha t} (1 + t(z - \alpha)) ,$$

so that

$$s_0(t) = (1 - \alpha t) e^{\alpha t}, \quad s_1(t) = t e^{\alpha t} .$$

Defining

$$s \equiv \frac{\alpha + \beta}{2} = \frac{\operatorname{tr} A}{2}, \quad q \equiv \frac{\alpha - \beta}{2} = \pm \sqrt{-\det(A - sI)},$$

we have

$$s_0(t) = e^{st} \left(\cosh(qt) - s \frac{\sinh(qt)}{q} \right), \quad s_1(t) = e^{st} \frac{\sinh(qt)}{q},$$

where $\sinh(qt)/q$ is 0 if $t = 0$, and t if $q = 0$. Thus,

$$e^{tA} = e^{st} \left(\left(\cosh(qt) - s \frac{\sinh(qt)}{q} \right) I + \frac{\sinh(qt)}{q} A \right) .$$

Thus, as indicated above, the matrix A having decomposed into the sum of two mutually commuting pieces, the traceful piece and the traceless piece,

$$A = sI + (A - sI) ,$$

the matrix exponential reduces to a plain product of the exponentials of the two respective pieces. This is a formula often used in physics, as it amounts to the analog of Euler's formula for Pauli spin matrices, that is rotations of the doublet representation of the group $SU(2)$.

The polynomial S_t can also be given the following "interpolation" characterization. Define $e_t(z) \equiv e^{tz}$, and $n \equiv \deg P$. Then $S_t(z)$ is the unique degree $< n$ polynomial which satisfies $S_t^{(k)}(a) = e_t^{(k)}(a)$ whenever k is less than the multiplicity of a as a root of P . We assume, as we obviously can, that P is the minimal polynomial of A . We further assume that A is a diagonalizable matrix. In particular, the roots of P are simple, and the "interpolation" characterization indicates that S_t is given by the Lagrange interpolation formula, so it is the Lagrange–Sylvester polynomial .

At the other extreme, if $P = (z-a)^n$, then

$$S_t = e^{at} \sum_{k=0}^{n-1} \frac{t^k}{k!} (z-a)^k .$$

The simplest case not covered by the above observations is when $P = (z-a)^2(z-b)$ with $a \neq b$, which yields

$$S_t = e^{at} \frac{z-b}{a-b} \left(1 + \left(t + \frac{1}{b-a} \right) (z-a) \right) + e^{bt} \frac{(z-a)^2}{(b-a)^2} .$$

Evaluation by implementation of Sylvester's formula

A practical, expedited computation of the above reduces to the following rapid steps. Recall from above that an $n \times n$ matrix $\exp(tA)$ amounts to a linear combination of the first $n-1$ powers of A by the Cayley-Hamilton theorem. For diagonalizable matrices, as illustrated above, e.g. in the 2×2 case, Sylvester's formula yields $\exp(tA) = B_\alpha \exp(t\alpha) + B_\beta \exp(t\beta)$, where the B s are the Frobenius covariants of A .

It is easiest, however, to simply solve for these B s directly, by evaluating this expression and its first derivative at $t=0$, in terms of A and I , to find the same answer as above.

But this simple procedure also works for defective matrices, in a generalization due to Buchheim.^[14] This is illustrated here for a 4×4 example of a matrix which is *not diagonalizable*, and the B s are not projection matrices.

Consider

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1/8 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix} ,$$

with eigenvalues $\lambda_1=3/4$ and $\lambda_2=1$, each with a multiplicity of two.

Consider the exponential of each eigenvalue multiplied by t , $\exp(\lambda_i t)$. Multiply each exponentiated eigenvalue by the corresponding undetermined coefficient matrix B_i . If the eigenvalues have an algebraic multiplicity greater than 1, then repeat the process, but now multiplying by an extra factor of t for each repetition, to ensure linear independence.

(If one eigenvalue had a multiplicity of three, then there would be the three terms: $B_{i_1} e^{\lambda_i t}$, $B_{i_2} t e^{\lambda_i t}$, $B_{i_3} t^2 e^{\lambda_i t}$. By contrast, when all eigenvalues are distinct, the B s are just the Frobenius covariants, and solving for them as below just amounts to the inversion of the Vandermonde matrix of these 4 eigenvalues.)

Sum all such terms, here four such,

$$\begin{aligned} e^{At} &= B_{1_1} e^{\lambda_1 t} + B_{1_2} t e^{\lambda_1 t} + B_{2_1} e^{\lambda_2 t} + B_{2_2} t e^{\lambda_2 t}, \\ e^{At} &= B_{1_1} e^{3/4t} + B_{1_2} t e^{3/4t} + B_{2_1} e^{1t} + B_{2_2} t e^{1t}. \end{aligned}$$

To solve for all of the unknown matrices B in terms of the first three powers of A and the identity, one needs four equations, the above one providing one such at $t=0$. Further, differentiate it with respect to t ,

$$A e^{At} = 3/4 B_{1_1} e^{3/4t} + (3/4t + 1) B_{1_2} e^{3/4t} + 1 B_{2_1} e^{1t} + (1t + 1) B_{2_2} e^{1t},$$

and again,

$$\begin{aligned} A^2 e^{At} &= (3/4)^2 B_{1_1} e^{3/4t} + ((3/4)^2 t + (3/4 + 1 \cdot 3/4)) B_{1_2} e^{3/4t} + B_{2_1} e^{1t} \\ &\quad + (1^2 t + (1 + 1 \cdot 1)) B_{2_2} e^{1t} \\ &= (3/4)^2 B_{1_1} e^{3/4t} + ((3/4)^2 t + 3/2) B_{1_2} e^{3/4t} + B_{2_1} e^{1t} + (t + 2) B_{2_2} e^{1t}, \end{aligned}$$

and once more,

$$\begin{aligned} A^3 e^{At} &= (3/4)^3 B_{1_1} e^{3/4t} + ((3/4)^3 t + ((3/4)^2 + (3/2) \cdot 3/4)) B_{1_2} e^{3/4t} \\ &\quad + B_{2_1} e^{1t} + (1^3 t + (1 + 2) \cdot 1) B_{2_2} e^{1t} \\ &= (3/4)^3 B_{1_1} e^{3/4t} + ((3/4)^3 t + 27/16) B_{1_2} e^{3/4t} + B_{2_1} e^{1t} + (t + 3 \cdot 1) B_{2_2} e^{1t} \end{aligned}$$

(In the general case, $n-1$ derivatives need be taken.)

Setting $t=0$ in these four equations, the four coefficient matrices B s may now be solved for,

$$\begin{aligned} I &= B_{1_1} + B_{2_1} \\ A &= 3/4 B_{1_1} + B_{1_2} + B_{2_1} + B_{2_2} \\ A^2 &= (3/4)^2 B_{1_1} + (3/2) B_{1_2} + B_{2_1} + 2 B_{2_2} \\ A^3 &= (3/4)^3 B_{1_1} + (27/16) B_{1_2} + B_{2_1} + 3 B_{2_2} \end{aligned},$$

to yield

$$\begin{aligned} B_{1_1} &= 128A^3 - 366A^2 + 288A - 80I \\ B_{1_2} &= 16A^3 - 44A^2 + 40A - 12I \\ B_{2_1} &= -128A^3 + 366A^2 - 288A + 80I \\ B_{2_2} &= 16A^3 - 40A^2 + 33A - 9I \end{aligned}$$

Substituting with the value for A yields the coefficient matrices

$$\begin{aligned}
 B_{1_1} &= \begin{pmatrix} 0 & 0 & 48 & -16 \\ 0 & 0 & -8 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
 B_{1_2} &= \begin{pmatrix} 0 & 0 & 4 & -2 \\ 0 & 0 & -1 & 1/2 \\ 0 & 0 & 1/4 & -1/8 \\ 0 & 0 & 1/2 & -1/4 \end{pmatrix} \\
 B_{2_1} &= \begin{pmatrix} 1 & 0 & -48 & 16 \\ 0 & 1 & 8 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 B_{2_2} &= \begin{pmatrix} 0 & 1 & 8 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

so the final answer is

$$e^{tA} = \begin{pmatrix} e^t & te^t & (8t - 48)e^t + (4t + 48)e^{3t/4} & (16 - 2t)e^t + (-2t - 16)e^{3t/4} \\ 0 & e^t & 8e^t + (-t - 8)e^{3t/4} & -\frac{4e^t + (-t - 4)e^{3t/4}}{2} \\ 0 & 0 & \frac{(t+4)e^{3t/4}}{4} & -\frac{te^{3t/4}}{8} \\ 0 & 0 & \frac{te^{3t/4}}{2} & -\frac{(t-4)e^{3t/4}}{4} \end{pmatrix}.$$

The procedure is much shorter than Putzer's algorithm sometimes utilized in such cases.

Illustrations

Suppose that we want to compute the exponential of

$$B = \begin{bmatrix} 21 & 17 & 6 \\ -5 & -1 & -6 \\ 4 & 4 & 16 \end{bmatrix}.$$

Its Jordan form is

$$J = P^{-1}BP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 16 & 1 \\ 0 & 0 & 16 \end{bmatrix},$$

where the matrix P is given by

$$P = \begin{bmatrix} -\frac{1}{4} & 2 & \frac{5}{4} \\ \frac{1}{4} & -2 & -\frac{1}{4} \\ 0 & 4 & 0 \end{bmatrix}.$$

Let us first calculate $\exp(J)$. We have

$$J = J_1(4) \oplus J_2(16)$$

The exponential of a 1×1 matrix is just the exponential of the one entry of the matrix, so $\exp(J_1(4)) = [e^4]$. The exponential of $J_2(16)$ can be calculated by the formula $e^{(\lambda I + N)} = e^\lambda e^N$ mentioned above; this yields^[15]

$$\begin{aligned} \exp\left(\begin{bmatrix} 16 & 1 \\ 0 & 16 \end{bmatrix}\right) &= e^{16} \exp\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) \\ &= e^{16} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \cdots \right) = \begin{bmatrix} e^{16} & e^{16} \\ 0 & e^{16} \end{bmatrix}. \end{aligned}$$

Therefore, the exponential of the original matrix B is

$$\begin{aligned} \exp(B) &= P \exp(J) P^{-1} = P \begin{bmatrix} e^4 & 0 & 0 \\ 0 & e^{16} & e^{16} \\ 0 & 0 & e^{16} \end{bmatrix} P^{-1} \\ &= \frac{1}{4} \begin{bmatrix} 13e^{16} - e^4 & 13e^{16} - 5e^4 & 2e^{16} - 2e^4 \\ -9e^{16} + e^4 & -9e^{16} + 5e^4 & -2e^{16} + 2e^4 \\ 16e^{16} & 16e^{16} & 4e^{16} \end{bmatrix}. \end{aligned}$$

Applications

Linear differential equations

The matrix exponential has applications to systems of linear differential equations. (See also matrix differential equation.) Recall from earlier in this article that a *homogeneous* differential equation of the form

$$\mathbf{y}' = A\mathbf{y}$$

has solution $e^{At} \mathbf{y}(0)$.

If we consider the vector

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix},$$

we can express a system of *inhomogeneous* coupled linear differential equations as

$$\mathbf{y}'(t) = A\mathbf{y}(t) + \mathbf{b}(t).$$

Making an ansatz to use an integrating factor of e^{-At} and multiplying throughout, yields

$$\begin{aligned} e^{-At} \mathbf{y}' - e^{-At} A \mathbf{y} &= e^{-At} \mathbf{b} \\ e^{-At} \mathbf{y}' - A e^{-At} \mathbf{y} &= e^{-At} \mathbf{b} \\ \frac{d}{dt}(e^{-At} \mathbf{y}) &= e^{-At} \mathbf{b}. \end{aligned}$$

The second step is possible due to the fact that, if $AB = BA$, then $e^{At}B = Be^{At}$. So, calculating e^{At} leads to the solution to the system, by simply integrating the third step with respect to t .

Example (homogeneous)

Consider the system

$$\begin{aligned} x' &= 2x - y + z \\ y' &= 3y - 1z \\ z' &= 2x + y + 3z. \end{aligned}$$

The associated defective matrix is

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}.$$

The matrix exponential is

$$e^{tA} = \frac{1}{2} \begin{bmatrix} e^{2t}(1 + e^{2t} - 2t) & -2te^{2t} & e^{2t}(-1 + e^{2t}) \\ -e^{2t}(-1 + e^{2t} - 2t) & 2(t+1)e^{2t} & -e^{2t}(-1 + e^{2t}) \\ e^{2t}(-1 + e^{2t} + 2t) & 2te^{2t} & e^{2t}(1 + e^{2t}) \end{bmatrix},$$

so that the general solution of the homogeneous system is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{x(0)}{2} \begin{bmatrix} e^{2t}(1 + e^{2t} - 2t) \\ -e^{2t}(-1 + e^{2t} - 2t) \\ e^{2t}(-1 + e^{2t} + 2t) \end{bmatrix} + \frac{y(0)}{2} \begin{bmatrix} -2te^{2t} \\ 2(t+1)e^{2t} \\ 2te^{2t} \end{bmatrix} + \frac{z(0)}{2} \begin{bmatrix} e^{2t}(-1 + e^{2t}) \\ -e^{2t}(-1 + e^{2t}) \\ e^{2t}(1 + e^{2t}) \end{bmatrix},$$

amounting to

$$\begin{aligned} 2x &= x(0)(e^{2t}(1 + e^{2t} - 2t)) + y(0)(-2te^{2t}) + z(0)(e^{2t}(-1 + e^{2t})) \\ 2y &= x(0)(-e^{2t}(-1 + e^{2t} - 2t)) + y(0)(2(t+1)e^{2t}) + z(0)(-e^{2t}(-1 + e^{2t})) \\ 2z &= x(0)(e^{2t}(-1 + e^{2t} + 2t)) + y(0)(2te^{2t}) + z(0)(e^{2t}(1 + e^{2t})). \end{aligned}$$

Example (inhomogeneous)

Consider now the inhomogeneous system

$$\begin{aligned}x' &= 2x - y + z + e^{2t} \\y' &= 3y - z \\z' &= 2x + y + 3z + e^{2t}.\end{aligned}$$

We again have

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix},$$

and

$$\mathbf{b} = e^{2t} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

From before, we already have the general solution to the homogeneous equation. Since the sum of the homogeneous and particular solutions give the general solution to the inhomogeneous problem, we now only need find the particular solution.

We have, by above,

$$\begin{aligned}\mathbf{y}_p &= e^{tA} \int_0^t e^{(-u)A} \begin{bmatrix} e^{2u} \\ 0 \\ e^{2u} \end{bmatrix} du + e^{tA} \mathbf{c} \\ \mathbf{y}_p &= e^{tA} \int_0^t \begin{bmatrix} 2e^u - 2ue^{2u} & -2ue^{2u} & 0 \\ -2e^u + 2(u+1)e^{2u} & 2(u+1)e^{2u} & 0 \\ 2ue^{2u} & 2ue^{2u} & 2e^u \end{bmatrix} \begin{bmatrix} e^{2u} \\ 0 \\ e^{2u} \end{bmatrix} du + e^{tA} \mathbf{c} \\ \mathbf{y}_p &= e^{tA} \int_0^t \begin{bmatrix} e^{2u}(2e^u - 2ue^{2u}) \\ e^{2u}(-2e^u + 2(1+u)e^{2u}) \\ 2e^{3u} + 2ue^{4u} \end{bmatrix} du + e^{tA} \mathbf{c} \\ \mathbf{y}_p &= e^{tA} \begin{bmatrix} -\frac{1}{24}e^{3t}(3e^t(4t-1)-16) \\ \frac{1}{24}e^{3t}(3e^t(4t+4)-16) \\ \frac{1}{24}e^{3t}(3e^t(4t-1)-16) \end{bmatrix} + \begin{bmatrix} 2e^t - 2te^{2t} & -2te^{2t} & 0 \\ -2e^t + 2(t+1)e^{2t} & 2(t+1)e^{2t} & 0 \\ 2te^{2t} & 2te^{2t} & 2e^t \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},\end{aligned}$$

which could be further simplified to get the requisite particular solution determined through variation of parameters. Note $\mathbf{c} = \mathbf{y}_p(0)$. For more rigor, see the following generalization.

Inhomogeneous case generalization: variation of parameters

For the inhomogeneous case, we can use integrating factors (a method akin to variation of parameters). We seek a particular solution of the form $\mathbf{y}_p(t) = \exp(tA) \mathbf{z}(t)$,

$$\begin{aligned}\mathbf{y}_p'(t) &= (e^{tA})' \mathbf{z}(t) + e^{tA} \mathbf{z}'(t) \\ &= A e^{tA} \mathbf{z}(t) + e^{tA} \mathbf{z}'(t) \\ &= A \mathbf{y}_p(t) + e^{tA} \mathbf{z}'(t) .\end{aligned}$$

For \mathbf{y}_p to be a solution,

$$\begin{aligned}e^{tA} \mathbf{z}'(t) &= \mathbf{b}(t) \\ \mathbf{z}'(t) &= (e^{tA})^{-1} \mathbf{b}(t) \\ \mathbf{z}(t) &= \int_0^t e^{-uA} \mathbf{b}(u) du + \mathbf{c} .\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{y}_p(t) &= e^{tA} \int_0^t e^{-uA} \mathbf{b}(u) du + e^{tA} \mathbf{c} \\ &= \int_0^t e^{(t-u)A} \mathbf{b}(u) du + e^{tA} \mathbf{c} ,\end{aligned}$$

where \mathbf{c} is determined by the initial conditions of the problem.

More precisely, consider the equation

$$\mathbf{Y}' - A \mathbf{Y} = \mathbf{F}(t)$$

with the initial condition $\mathbf{Y}(t_0) = \mathbf{Y}_0$, where A is an n by n complex matrix,

\mathbf{F} is a continuous function from some open interval I to \mathbb{C}^n ,

t_0 is a point of I , and

\mathbf{Y}_0 is a vector of \mathbb{C}^n .

Left-multiplying the above displayed equality by e^{-tA} yields

$$\mathbf{Y}(t) = e^{(t-t_0)A} \mathbf{Y}_0 + \int_{t_0}^t e^{(t-x)A} \mathbf{F}(x) dx .$$

We claim that the solution to the equation

$$P(d/dt) \mathbf{y} = \mathbf{f}(t)$$

with the initial conditions $\mathbf{y}^{(k)}(t_0) = \mathbf{y}_k$ for $0 \leq k < n$ is

$$y(t) = \sum_{k=0}^{n-1} y_k s_k(t - t_0) + \int_{t_0}^t s_{n-1}(t - x) f(x) dx ,$$

where the notation is as follows:

$P \in \mathbb{C}[X]$ is a monic polynomial of degree $n > 0$,

f is a continuous complex valued function defined on some open interval I ,

t_0 is a point of I ,

y_k is a complex number, and

$s_k(t)$ is the coefficient of X^k in the polynomial denoted by $S_t \in \mathbb{C}[X]$ in Subsection Evaluation by Laurent series above.

To justify this claim, we transform our order n scalar equation into an order one vector equation by the usual reduction to a first order system. Our vector equation takes the form

$$\frac{dY}{dt} - A Y = F(t), \quad Y(t_0) = Y_0,$$

where A is the transpose companion matrix of P . We solve this equation as explained above, computing the matrix exponentials by the observation made in Subsection Evaluation by implementation of Sylvester's formula above.

In the case $n = 2$ we get the following statement. The solution to

$$y'' - (\alpha + \beta) y' + \alpha \beta y = f(t), \quad y(t_0) = y_0, \quad y'(t_0) = y_1$$

is

$$y(t) = y_0 s_0(t - t_0) + y_1 s_1(t - t_0) + \int_{t_0}^t s_1(t - x) f(x) dx,$$

where the functions s_0 and s_1 are as in Subsection Evaluation by Laurent series above.

Matrix-matrix exponentials

The matrix exponential of another matrix (matrix-matrix exponential),^[16] is defined as

$$\begin{aligned} X^Y &= e^{\log(X) \cdot Y} \\ {}^Y X &= e^{Y \cdot \log(X)} \end{aligned}$$

for X any normal and non-singular $n \times n$ matrix, and Y any complex $n \times n$ matrix.

For matrix-matrix exponentials, there is a distinction between the left exponential ${}^Y X$ and the right exponential X^Y , because the multiplication operator for matrix-to-matrix is not commutative. Moreover,

- If X is normal and non-singular, then X^Y and ${}^Y X$ have the same set of eigenvalues.
- If X is normal and non-singular, Y is normal, and $XY = YX$, then $X^Y = {}^Y X$.

- If X is normal and non-singular, and X, Y, Z commute with each other, then $X^{Y+Z} = X^Y \cdot X^Z$ and $X^{Y+Z} = X^Y \cdot X^Z$.

See also

- Matrix function
- Matrix logarithm
- Exponential function
- Exponential map (Lie theory)
- Magnus expansion
- Derivative of the exponential map
- Vector flow
- Golden–Thompson inequality
- Phase-type distribution
- Lie product formula
- Baker–Campbell–Hausdorff formula
- Frobenius covariant
- Sylvester's formula
- Trigonometric functions of matrices

References

1. Hall 2015 Equation 2.1
 2. Hall 2015 Proposition 2.3
 3. Bhatia, R. (1997). *Matrix Analysis*. Graduate Texts in Mathematics. **169**. Springer. ISBN 978-0-387-94846-1.
 4. E. H. Lieb (1973). "Convex trace functions and the Wigner–Yanase–Dyson conjecture". *Adv. Math.* **11** (3): 267–288. doi:10.1016/0001-8708(73)90011-X (https://doi.org/10.1016%2F0001-8708%2873%2990011-X).
 - H. Epstein (1973). "Remarks on two theorems of E. Lieb". *Commun Math. Phys.* **31** (4): 317–325. doi:10.1007/BF01646492 (https://doi.org/10.1007%2F01646492).
 5. Hall 2015 Exercises 2.9 and 2.10
 6. R. M. Wilcox (1967). "Exponential Operators and Parameter Differentiation in Quantum Physics". *Journal of Mathematical Physics*. **8** (4): 962–982. doi:10.1063/1.1705306 (https://doi.org/10.1063%2F1.1705306).
 7. Hall 2015 Theorem 2.12
 8. "Matrix exponential - MATLAB expm - MathWorks Deutschland" (http://www.mathworks.de/help/techdoc/ref/expm.html). Mathworks.de. 2011-04-30. Retrieved 2013-06-05.
 9. "GNU Octave - Functions of a Matrix" (http://www.network-theory.co.uk/docs/octave3/octave_200.html). Network-theory.co.uk. 2007-01-11. Retrieved 2013-06-05.
 10. "scipy.linalg.expm function documentation" (http://docs.scipy.org/doc/scipy-0.15.1/reference/generated/scipy.linalg.expm.html). The SciPy Community. 2015-01-18. Retrieved 2015-05-29.
 11. in a Euclidean space
 12. Weyl, Hermann (1952). *Space Time Matter* (https://books.google.com/books?id=KCgZAQAIAAJ&pg=PA142#v=onepage&q&f=false). Dover. p. 142. ISBN 0-486-60267-2.
 13. Bjorken, James D.; Drell, Sidney D. (1964). *Relativistic Quantum Mechanics*. McGraw-Hill. p. 22.
 14. Rinehart, R. F. (1955). "The equivalence of definitions of a matrix function". *The American Mathematical Monthly*, **62** (6), 395-414.
 15. This can be generalized; in general, the exponential of $J_n(a)$ is an upper triangular matrix with $e^a/0!$ on the main diagonal, $e^a/1!$ on the one above, $e^a/2!$ on the next one, and so on.
 16. Ignacio Barradas and Joel E. Cohen (1994). "Iterated Exponentiation, Matrix-Matrix Exponentiation, and Entropy" (http://www.rockefeller.edu/labheads/cohenje/PDFs/215BarrabasCohenalApp19941.pdf) (PDF). Academic Press, Inc.
- Hall, Brian C. (2015), *Lie groups, Lie algebras, and representations: An elementary introduction*, Graduate Texts in Mathematics, **222** (2nd ed.), Springer
 - Horn, Roger A.; Johnson, Charles R. (1991). *Topics in Matrix Analysis*. Cambridge University Press. ISBN 978-0-521-46713-1..

- Moler, Cleve; Van Loan, Charles F. (2003). "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later" (<http://www.cs.cornell.edu/cv/researchpdf/19ways+.pdf>) (PDF). *SIAM Review*. **45** (1): 3–49. ISSN 1095-7200 (<https://www.worldcat.org/issn/1095-7200>). doi:10.1137/S00361445024180 (<https://doi.org/10.1137%2FS00361445024180>)..
- Suzuki, Masuo (1985). "Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics". *Journal of Mathematical Physics*. **26**: 601. doi:10.1063/1.526596 (<https://doi.org/10.1063%2F1.526596>).
- Curtright, T L; Fairlie, D B; Zachos, C K (2014). "A compact formula for rotations as spin matrix polynomials". *SIGMA*. **10**: 084. doi:10.3842/SIGMA.2014.084 (<https://doi.org/10.3842%2FSIGMA.2014.084>).
- Householder, Alston S. (2006). *The Theory of Matrices in Numerical Analysis*. Dover Books on Mathematics. ISBN 0486449726.
- Van Kortryk, T. S. (2016). "Matrix exponentials, SU(N) group elements, and real polynomial roots". *Journal of Mathematical Physics*. **57**: 021701. doi:10.1063/1.4938418 (<https://doi.org/10.1063%2F1.4938418>).

External links

- Weisstein, Eric W. "Matrix Exponential" (<http://mathworld.wolfram.com/MatrixExponential.html>). *MathWorld*.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Matrix_exponential&oldid=801794906"

-
- This page was last edited on 21 September 2017, at 22:18.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.