4.X matrix exponential

Monday, March 15, 2021 2:21 PM

Back to basics: What is e^x? $(x \in \mathbb{R})$

$$(2.71828\ldots)^x$$

$\uparrow$ Jacob Bernoulli (1683)

$$e = \lim_{m \to \infty} \left(1 + \frac{1}{m}\right)^m$$ (studying compound interest)

$$(2.71828\ldots)^3 = (2.71828\ldots)(2.7182\ldots)(2.71828\ldots)$$

$$(2.71828\ldots)^{3.1} = (2.71828\ldots)^3 \sqrt[10]{2.71828\ldots}$$

$$(2.71828\ldots)^{3.14} = (2.71828\ldots)^3 \sqrt[10]{2.71828\ldots}\left(\sqrt[100]{2.71828\ldots}\right)^4$$

$$e^{\pi} = (2.71828\ldots)^{3.14\ldots} = \lim_{m \to \infty} (2.71828\ldots)^{x_m}, \quad \text{where } \lim_{m \to \infty} x_m = \pi$$

Def. 1: e^x is the unique solution to $\frac{d}{dx}[y(x)] = y(x)$, $y(0) = 1$.

Def. 2: $e^x = \lim_{m \to \infty} \left(1 + \frac{x}{m}\right)^m$

Def. 3: $e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!}$

Note: Since $e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!}$ has infinite radius of convergence, so we can use this definition for all $e^z = \sum_{m=0}^{\infty} \frac{z^m}{m!}$, where $z \in \mathbb{C}$.

Matrix exponential:

Def. 3: Let $e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!}$, where $A \in \mathbb{C}^{n \times n}$

Properties: (1) $e^0 = I$, where $0 \in \mathbb{C}^{n \times n}$ is 0 matrix.

(2) [illegible] $e^X e^Y = e^{X+Y}$

Properties: (1) $e^{0} = I$, where $0 \in \mathbb{C}$ is 0 matrix.

(2) If $AB = BA$, $A, B \in \mathbb{C}^{n\times n}$, then $e^X e^Y = e^{X+Y}$

(3) If $P \in \mathbb{C}^{n\times n}$ is invertible, then $e^{PBP^{-1}} = P e^B P^{-1}$

Suppose $A = PBP^{-1}$

Then $A^m = (PBP^{-1})^m = \underbrace{(PB\overbrace{P^{-1})(P}^{I}BP^{-1}) \cdots (PBP^{-1})}_{m \text{ times}} = P B^m P^{-1}$

Thus $e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!} = \sum_{m=0}^{\infty} \frac{1}{m!} \cdot P B^m P^{-1} = P\left(\sum_{m=0}^{\infty} \frac{1}{m} \cdot B^m\right) P^{-1} = P\left(e^B\right)P^{-1}$.

Diagonalizable matrix $A = PDP^{-1}$, where D is diagonal.

Let
$$D = \begin{bmatrix} d_{11} & & & 0 \\ & d_{22} & & \\ & & \ddots & \\ 0 & & & d_{nn} \end{bmatrix}.$$
Then
$$D^m = \begin{bmatrix} d_{11}^m & & 0 \\ & \ddots & \\ 0 & & d_{nn}^m \end{bmatrix}$$

So
$$e^D = \sum_{m=0}^{\infty} \frac{D^m}{m!} = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{d_{11}^m}{m!} & & \\ & \ddots & \\ & & \sum_{m=0}^{\infty} \frac{d_{nn}^m}{m!} \end{bmatrix} = \begin{bmatrix} e^{d_{11}} & & \\ & \ddots & \\ & & e^{d_{nn}} \end{bmatrix}$$

$$\Rightarrow \quad e^A = P \begin{bmatrix} e^{d_{11}} & & 0 \\ & \ddots & \\ 0 & & e^{d_{nn}} \end{bmatrix} P^{-1}, \text{ where } A = PDP^{-1}.$$

Non-diagonalizable A

Recall Jordan Canonical Form:

Every square matrix $A \in \mathbb{C}^{n\times n}$ can be transformed into

$\ldots = \begin{bmatrix} J_1 & & 0 \\ \ldots \end{bmatrix}$

Every square matrix $A \in \mathbb{C}^{n\times n}$ can be transformed into

$$A = PJP^{-1}, \quad \text{where } P \in \mathbb{C}^{n\times n} \text{ and } J = \underbrace{\begin{bmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_p \end{bmatrix}}_{\text{block diagonal}}$$

Each J_i is a Jordan block with λ_i on the diagonal, 1 directly above the diagonal, and 0 everywhere else.

$$J_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}$$

Then $e^A = Pe^J P^{-1}$ by property 3 above.

But

$$e^J = \sum_{m=0}^{\infty} \frac{J^m}{m!} = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{J_1^m}{m!} & & \\ & \ddots & \\ & & \sum_{m=0}^{\infty} \frac{J_p^m}{m!} \end{bmatrix} = \begin{bmatrix} e^{J_1} & & \\ & \ddots & \\ & & e^{J_p} \end{bmatrix}$$

$$\exp(J_i) = \exp\left(\begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}\right) = \exp\left(\begin{bmatrix} \lambda_i & & 0 \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix} + \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}\right)$$

$$= \exp(\lambda_i I)\exp\left(\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}\right) \qquad \text{(Prop. 2)}$$

Let's call $\overset{\text{nilpotent}}{N_q} \in \mathbb{C}^{q\times q}$, the $q\times q$ matrix with 1's in the row above the diagonal and 0's everywhere else,

Ex. $$N_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad N_4^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then N_q^2 has 1's two rows above the diagonal and 0's everywhere else,

Then N_q^k has 1's k rows above the diagonal and 0's everywhere else

$\Rightarrow N_q^q = 0.$

Thus, $\exp(N_q) = \sum_{m=0}^{\infty} \frac{N_q^m}{m!} = \sum_{m=0}^{q-1} \frac{N_q^m}{m!} = I + N_q + \frac{1}{2!}N_q^2 + \frac{1}{3!}N_q^3 + \cdots + \frac{1}{(q-1)!}N_q^{q-1}.$

$$= \begin{bmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{6} & \cdots & \frac{1}{(q-1)!} \\ & 1 & 1 & \frac{1}{2} & \frac{1}{6} & \vdots \\ & & \ddots & \ddots & \ddots & \frac{1}{6} \\ & & & \ddots & \ddots & \frac{1}{2} \\ & 0 & & & \ddots & 1 \\ & & & & & 1 \end{bmatrix}$$

Thus, we can compute

$\exp(J_i) = \exp(\lambda_i I)\exp(N_{q_i})$, where q_i is the size of the ith Jordan block.

And $\exp(A) = P(\exp(J))P^{-1}$ is computable.

Computing matrix exponentials using differential equations:

Recall: Let $\frac{dX}{dt} = AX$, $A \in \mathbb{R}^{n\times n}$, $X = X(t) \in \mathbb{R}^n$

Then $X(t) = X(0)\exp(At)$ ← Picard iteration proof

Could expand $\exp(At) = \sum_{m=0}^{\infty} A^m \frac{t^m}{m!}$ to solve $X(t)$.

On the other hand, if we knew $X(t)$, can solve for $\exp(At)$.

[Leonard, 1996]: Let $A \in \mathbb{R}^{n\times n}$ matrix and its characteristic equation

[Leonard, 1996 SIAM Review]: Let $A \in \mathbb{R}^{n\times n}$ matrix and its characteristic equation

$$\det(\lambda I - A) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n = 0.$$

We can write a higher-order ODE with constant coefficients

$$\frac{d^n}{dt^n} x(t) + a_1 \frac{d^{n-1}}{dt^{n-1}} x(t) + \cdots + a_{n-1} \frac{d}{dt} x(t) + a_n x(t) = 0$$

Then, we can find n linearly ind. solutions $x_1(t), \dots, x_n(t)$ with initial conditions

$$\boxed{\begin{matrix} x_1(0)=1 \\ \frac{d}{dt}x_1(0)=0 \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}} x_1(0) = 0 \end{matrix}}, \quad \boxed{\begin{matrix} x_2(0)=0 \\ \frac{d}{dt}x_2(0)=1 \\ \frac{d^2}{dt^2}x_2(0)=0 \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}} x_2(0) = 0 \end{matrix}}, \quad \dots, \quad \boxed{\begin{matrix} x_n(0)=0 \\ \frac{d}{dt}x_n(0)=0 \\ \vdots \\ \frac{d^{n-2}}{dt^{n-2}} x_n(0) = 0 \\ \frac{d^{n-1}}{dt^{n-1}} x_n(0) = 1 \end{matrix}}$$

Then $e^{At} = x_1(t) I + x_2(t) A + \cdots + x_n(t) A^{n-1}$.

Exercise for viewer: Verify that the alternate definition

$$e^A = \lim_{m\to\infty} \left(I + \frac{1}{m} A\right)^m$$ also works.