5. Function of square matrices

- matrix polynomial
- rational function
- Cayley-Hamilton theotem
- infinite series
- matrix exponential
- applications to differential equations

Matrix Power

the mth power of a matrix A for a *nonnegative* m is defined as

$$A^m = \prod_{k=1}^{m} A$$

and define $A^0 = I$

property: $A^r A^s = A^s A^r = A^{r+s}$

a *negative* power of A is defined as

$$A^{-n} = (A^{-1})^n$$

n is a nonnegative integer and A is invertible

Matrix polynomial

a **matrix polynomial** is a polynomial with matrices as variables

$$p(A) = a_0 I + a_1 A + \cdots + a_n A^n$$

for example $A = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$

$$\begin{aligned} p(A) = 2I - 6A + 3A^2 \quad &= \quad 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 6\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix} + 3\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}^2 \\ &= \quad \begin{bmatrix} 2 & -3 \\ 0 & 11 \end{bmatrix} \end{aligned}$$

Fact ✎ any two polynomials of A commute, *i.e.*, $p(A)q(A) = q(A)p(A)$

similarity transform: suppose A is diagonalizable, *i.e.*,

$$\Lambda = T^{-1}AT \quad \Longleftrightarrow \quad A = T\Lambda T^{-1}$$

where $T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$, *i.e.*, the columns of T are eigenvectors of A

then we have $A^k = T\Lambda^k T^{-1}$

thus diagonalization simplifies the expression of a matrix polynomial

$$\begin{aligned} p(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= a_0 TT^{-1} + a_1 T\Lambda T^{-1} + \cdots + a_n T\Lambda^n T^{-1} \\ &= Tp(\Lambda)T^{-1} \end{aligned}$$

where

$$p(\Lambda) = \begin{bmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{bmatrix}$$

eigenvalues and eigenvectors ✌

if λ and v be an eigenvalue and corresponding eigenvector of A then

- $p(\lambda)$ is an eigenvalue of $p(A)$
- v is a corresponding eigenvector of $p(A)$

$$Av = \lambda v \quad \Longrightarrow \quad A^2 v = \lambda A v = \lambda^2 v \quad \cdots \quad \Longrightarrow \quad A^k v = \lambda^k v$$

thus

$$(a_0 I + a_1 A + \cdots + a_n A^n)v = (a_0 v + a_1 \lambda + \cdots + a_n \lambda^n)v$$

which shows that

$$p(A)v = p(\lambda)v$$

Rational functions

$f(x)$ is called a **rational function** if and only if it can be written as

$$f(x) = \frac{p(x)}{q(x)}$$

where $p(x)$ and $q(x)$ are polynomial functions in x and $q(x) \neq 0$

we define a rational function for square matrices as

$$f(A) = \frac{p(A)}{q(A)} \quad \triangleq \quad p(A)q(A)^{-1} = q^{-1}(A)p(A)$$

provided that $q(A)$ is invertible

eigenvalues and eigenvectors ✌

if λ and v be an eigenvalue and corresponding eigenvector of A then

- $p(\lambda)/q(\lambda)$ is an eigenvalue of $f(A)$
- v is a corresponding eigenvector of $f(A)$

both $p(A)$ and $q(A)$ are polynomials, so we have

$$p(A)v = p(\lambda)v, \quad q(A)v = q(\lambda)v$$

and the eigenvalue of $q(A)^{-1}$ is $1/q(\lambda)$, *i.e.*, $q(A)^{-1}v = (1/q(\lambda))v$

thus

$$f(A)v = p(A)q(A)^{-1}v = q(\lambda)^{-1}p(A)v = q(\lambda)^{-1}p(\lambda)v = f(\lambda)v$$

which says that $f(\lambda)$ is an eigenvalue of $f(A)$ with the same eigenvector

example: $f(x) = (x+1)/(x-5)$ and $A = \begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}$

$$\det(\lambda I - A) = 0 = (\lambda - 4)(\lambda - 5) - 2 = \lambda^2 - 9\lambda + 18 = 0$$

the eigenvalues of A are $\lambda_1 = 3$ and $\lambda_2 = 6$

$$f(A) = (A+I)(A-5I)^{-1} = \begin{bmatrix} 5 & 2 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 6 \\ 3 & 4 \end{bmatrix}$$

the characteristic function of $f(A)$ is

$$\det(\lambda I - f(A)) = 0 = (\lambda - 1)(\lambda - 4) - 18 = \lambda^2 - 5\lambda - 14 = 0$$

the eigenvalues of $f(A)$ are 7 and -2

this agrees to the fact that the eigenvalues of $f(A)$ are

$$f(\lambda_1) = (\lambda_1 - 1)/(\lambda_1 - 5) = -2, \quad f(\lambda_2) = (\lambda_2 - 1)/(\lambda_2 - 5) = 7$$

Cayley-Hamilton theorem

the characteristic polynomial of a matrix A of size $n \times n$

$$\mathcal{X}(\lambda) = \det(\lambda I - A)$$

can be written as a polynomial of degree n:

$$\mathcal{X}(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1\lambda + \alpha_0$$

✌ **Theorem:** a square matrix satisfies its characteristic equation:

$$\mathcal{X}(A) = A^n + \alpha_{n-1}A^{n-1} + \cdots + \alpha_1 A + \alpha_0 I = 0$$

result: for $m \geq n$, A^m is a linear combination of $A^k, k = 0, 1, \ldots, n-1$.

example 1: $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ the characteristic equation of A is

$$\mathcal{X}(\lambda) = (\lambda - 1)(\lambda - 3) = \lambda^2 - 4\lambda + 3 = 0$$

the Cayley-Hamilton states that A *satisfies its characteristic equation*

$$\mathcal{X}(A) = A^2 - 4A + 3I = 0$$

use this equation to write matrix powers of A

$$\begin{array}{rclclcl} A^2 & = & 4A - 3I & & & & \\ A^3 & = & 4A^2 - 3A & = & 4(4A - 3I) - 3A & = & 13A - 12I \\ A^4 & = & 13A^2 - 12A & = & 13(4A - 3I) - 12A & = & 40A - 39I \\ \vdots & & & & \vdots & & \end{array}$$

powers of A can be written as a linear combination of I and A

example 2: with A in page 5-10, find the closed-form expression of A^k

for $k \geq 2$, A^k is a linear combination of I and A, *i.e.*,

$$A^k = \alpha_0 I + \alpha_1 A$$

where α_1, α_0 are to be determined

multiply eigenvectors of A on both sides

$$\begin{aligned} A^k v_1 &= (\alpha_0 I + \alpha_1 A)v_1 \quad \Rightarrow \quad \lambda_1^k = \alpha_0 + \alpha_1 \lambda_1 \\ A^k v_2 &= (\alpha_0 I + \alpha_1 A)v_2 \quad \Rightarrow \quad \lambda_2^k = \alpha_0 + \alpha_1 \lambda_2 \end{aligned}$$

substitute $\lambda_1 = 1$ and $\lambda_2 = 3$ and solve for α_0, α_1

$$\begin{bmatrix} 1 \\ 3^k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \quad \Rightarrow \quad \alpha_0 = \frac{3 - 3^k}{2}, \quad \alpha_1 = \frac{3^k - 1}{2}$$

$$A^k = \frac{3 - 3^k}{2} I + \frac{3^k - 1}{2} A = \begin{bmatrix} 1 & 3^k - 1 \\ 0 & 3^k \end{bmatrix}, \quad k \geq 2$$

Computing the inverse of a matrix

A is a square matrix with the characteristic equation

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$$

by the C-H theorem, A satisfies the characteristic equation

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0$$

if A is invertible, multiply A^{-1} on both sides

$$A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I + a_0 A^{-1} = 0$$

thus the inverse of A can be alternatively computed by

$$A^{-1} = -\frac{1}{a_0}\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right)$$

example: given $A = \begin{bmatrix} 2 & -4 & -4 \\ 1 & -4 & -5 \\ 1 & 4 & 5 \end{bmatrix}$ find A^{-1}

the characteristic equation of A is

$$\det(\lambda I - A) = \lambda^3 - 3\lambda^2 + 10\lambda - 8 = 0$$

0 is not an eigenvalue of A, so A is invertible and given by

$$\begin{aligned} A^{-1} &= \frac{1}{8}\left(A^2 - 3A + 10I\right) \\ &= \frac{1}{4}\begin{bmatrix} 0 & 2 & 2 \\ -5 & 7 & 3 \\ 4 & -6 & 2 \end{bmatrix} \end{aligned}$$

compare the result with other methods

Infinite series

Definition: a series $\sum_{k=0}^{\infty} a_k$ *converges* to S if the partial sum

$$S_n \triangleq \sum_{k=0}^{n} a_k$$

converges to S as $n \to \infty$

example of convergent series:

$$\begin{aligned} 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots &= 2 \\ 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots &= \log(2) \end{aligned}$$

Power series

a power series in scalar variable z is an infinite series of the form

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

example: power series that converges for *all values* of z

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

$$\cos(z) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \cdots$$

$$\sin(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots$$

$$\cosh(z) = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \frac{z^6}{6!} + \cdots$$

$$\sinh(z) = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \frac{z^7}{7!} + \cdots$$

Power series of matrices

let A be matrix and A_{ij} denotes the (i,j) entry of A

Definition: a matrix power series

$$\sum_{k=0}^{\infty} a_k A_k$$

converges to S if all (i,j) entries of the partial sum

$$S_n \triangleq \sum_{k=0}^{n} a_k A_k$$

converges to the corresponding (i,j) entries of S as $n \to \infty$

Fact ✌ if $f(z) = \sum_{k=0}^{\infty} a_k z^k$ is a convergent power series for *all* z then

$f(A)$ is convergent for *any square matrix* A

Matrix exponential

generalize the exponential function of a scalar

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

to an exponential function of a matrix

define **matrix exponential** as

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots \quad = \quad \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

for a square matrix A

the infinite series converges for all A

example: $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

find all powers of A

$$A^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \ldots, \quad A^k = A \quad \text{for } k = 2, 3, \ldots$$

so by definition,

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + \sum_{k=1}^{\infty} \frac{A^k}{k!} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sum_{k=1}^{\infty} \frac{1}{k!} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}$$

never compute e^A by element-wise operation !

$$e^A \neq \begin{bmatrix} e^1 & e^1 \\ e^0 & e^0 \end{bmatrix}$$

Eigenvalues of matrix exponential

✌ if λ and v be an eigenvalue and corresponding eigenvector of A then

- e^{λ} is an eigenvalue of e^{A}
- v is a corresponding eigenvector of e^{A}

since e^{A} can be expressed as power series of A:

$$e^{A} = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

multiplying v on both sides and using $A^k v = \lambda^k v$ give

$$\begin{aligned} e^{A}v &= v + Av + \tfrac{A^2 v}{2!} + \tfrac{A^3 v}{3!} + \cdots \\ &= \left(1 + \lambda + \tfrac{\lambda^2}{2!} + \tfrac{\lambda^3}{3!} + \cdots\right) v \\ &= e^{\lambda} v \end{aligned}$$

Properties of matrix exponential

- $e^0 = I$
- $e^{A+B} \neq e^A \cdot e^B$
- if $AB = BA$, $i.e.$, A and B commute, then $e^{A+B} = e^A \cdot e^B$
- $(e^A)^{-1} = e^{-A}$

✌ these properties can be proved by the definition of e^A

Computing e^A via diagonalization

if A is diagonalizable, *i.e.*,

$$T^{-1}AT = \Lambda = \mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

where λ_k's are eigenvalues of A then e^A has the form

$$e^A = Te^{\Lambda}T^{-1}$$

- computing e^{Λ} is simple since Λ is diagonal
- one needs to find eigenvectors of A to form the matrix T
- the expression of e^A follows from

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{(T\Lambda T^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{T\Lambda^k T^{-1}}{k!} = Te^{\Lambda}T^{-1}$$

- if A is diagonalizable, so is e^A

example: compute $f(A) = e^A$ given $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

the eigenvalues and eigenvectors of A are

$$\lambda_1 = 1,\ v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \lambda_2 = 2,\ v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \lambda_3 = 0,\ v_3 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

form $T = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$ and compute $e^A = Te^\Lambda T^{-1}$

$$e^A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} e & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1/2 \end{bmatrix} = \begin{bmatrix} e & e^2 - e & (e^2 - 2e + 1)/2 \\ 0 & e^2 & (e^2 - 1)/2 \\ 0 & 0 & 1 \end{bmatrix}$$

Computing e^A via C-H theorem

e^A is an infinite series

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

by C-H theorem, the power A^k can be written as

$$A^k = a_0 I + a_1 A + \cdots + a_{n-1} A^{n-1}, \quad k = n, n+1, \ldots$$

(a polynomial in A of order $\leq n-1$)

thus e^A can be expressed as a linear combination of $I, A, \ldots, A^{n-1}$

$$e^A = \alpha_0 I + \alpha_1 A + \cdots + \alpha_{n-1} A^{n-1}$$

where α_k's are coefficients to be determined this also holds for any convergent power series $f(A) = \sum_{k=0}^{\infty} a_k A^k$

$$f(A) = \alpha_0 I + \alpha_1 A + \cdots + \alpha_{n-1} A^{n-1}$$

(recursively write A^k as a linear combination of $I, A, \ldots, A^{n-1}$ for $k \geq n$)

multiplying an eigenvector v of A on both sides and using $v \neq 0$, we get

$$f(\lambda) = \alpha_0 I + \alpha_1 \lambda + \cdots + \alpha_{n-1} \lambda^{n-1}$$

substitute with the n eigenvalues of A

$$\begin{bmatrix} f(\lambda_1) \\ f(\lambda_2) \\ \vdots \\ f(\lambda_n) \end{bmatrix} = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{bmatrix}$$

Fact ✎ if all λ_k's are distinct, the system is solvable and has a unique sol.

Vandermonde matrix

a *Vandermonde matrix* has the form

$$V = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

(with a geometric progression in each row)

✎ one can show that the determinant of V can be expressed as

$$\det(V) = \prod_{1 \leq i < j \leq n} (\lambda_j - \lambda_i)$$

hence, V is invertible as long as λ_i's are *distinct*

example: compute $f(A) = e^A$ given

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

the eigenvalues of A are $\lambda = 1, 2, 0$ (all are distinct)

form a system of equations: $f(\lambda_i) = \alpha_0 + \alpha_1 \lambda_i + \alpha_2 \lambda_i^2$ for $i = 1, 2, 3$

$$\begin{bmatrix} e^1 \\ e^2 \\ e^0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2^2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$$

which has the solution

$$\alpha_0 = 1, \quad \alpha_1 = 2e - e^2/2 - 3/2, \quad \alpha_2 = -e + e^2/2 + 1/2$$

substituting $\alpha_0, \alpha_1, \alpha_2$ in

$$e^A = \alpha_0 I + \alpha_1 A + \alpha_2 A^2$$

gives

$$\begin{aligned} e^A &= \alpha_0 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 & 3 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_0 + \alpha_1 + \alpha_2 & \alpha_1 + 3\alpha_2 & \alpha_2 \\ 0 & \alpha_0 + 2\alpha_1 + 4\alpha_2 & \alpha_1 + 2\alpha_2 \\ 0 & 0 & \alpha_0 \end{bmatrix} \\ &= \begin{bmatrix} e & e^2 - e & (e^2 - 2e + 1)/2 \\ 0 & e^2 & (e^2 - 1)/2 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

(agree with the result in page 5-22)

Repeated eigenvalues

A has repeated eigenvalues, *i.e.*, $\lambda_i = \lambda_j$ for some i, j

goal: compute $f(A)$ using C-H theorem

however, we can no longer apply the result in page 5-24 because

- the number of independent equations on page 5-24 is less than n
- the Vandermonde matrix (page 5-25) is not invertible

cannot form a linear system to solve for the n coefficients, $\alpha_0, \ldots, \alpha_{n-1}$

solution: for the repeated root with multiplicity r

get $r-1$ independent equations by taking derivatives on $f(\lambda)$ w.r.t λ

$$
\begin{aligned}
f(\lambda) &= \alpha_0 + \alpha_1\lambda + \cdots + \alpha_{n-1}\lambda^{n-1} \\
\frac{df(\lambda)}{d\lambda} &= \alpha_1 + 2\alpha_2\lambda + \cdots + (n-1)\alpha_{n-1}\lambda^{n-2} \\
\vdots &= \vdots \\
\frac{d^{r-1}f(\lambda)}{d^{r-1}\lambda} &= (r-1)!\alpha_{r-1} + \cdots + (n-r)\cdots(n-2)(n-1)\alpha_{n-1}\lambda^{n-1-r}
\end{aligned}
$$

example: compute $f(A) = \cos(A)$ given

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

the eigenvalues of A are $\lambda_1 = 1, 1$ and $\lambda_2 = 2$

by C-H theorem, write $f(A)$ as a linear combination of A^k, $k = 0, ..., n-1$

$$f(A) = \cos(A) = \alpha_0 + \alpha_1 A + \alpha_2 A^2$$

the eigenvalues of A must also satisfies this equation

$$f(\lambda) = \cos(\lambda) = \alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2$$

the derivative of f w.r.t λ is given by

$$f'(\lambda) = -\sin(\lambda) = \alpha_1 + 2\alpha_2 \lambda$$

thus we can obtain n linearly independent equations:

$$\begin{bmatrix} f(\lambda_1) \\ f'(\lambda_1) \\ f(\lambda_2) \end{bmatrix} = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 0 & 1 & 2\lambda_1 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} \Longrightarrow \begin{bmatrix} \cos(1) \\ -\sin(1) \\ \cos(2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$$

which have the solution

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 2\sin(1) + \cos(2) \\ 2\cos(1) - 3\sin(1) - 2\cos(2) \\ -\cos(1) + \sin(1) + \cos(2) \end{bmatrix}$$

substitute $\alpha_0, \alpha_1, \alpha_2$ to obtain $f(A)$

$$\begin{aligned} f(A) = \cos(A) &= \alpha_0 I + \alpha_1 A + \alpha_2 A^2 \\ &= \begin{bmatrix} \cos(1) & -\sin(1) & 0 \\ 0 & \cos(1) & 0 \\ 0 & 0 & \cos(2) \end{bmatrix} \end{aligned}$$

Applications to ordinary differential equations

we solve the following first-order ODEs for $t \geq 0$ where $x(0)$ is given

scalar: $x(t) \in \mathbf{R}$ and $a \in \mathbf{R}$ is given

$$\dot{x}(t) = ax(t)$$

solution: $x(t) = e^{at}x(0)$, for $t \geq 0$

vector: $x(t) \in \mathbf{R}^n$ and $A \in \mathbf{R}^{n \times n}$ is given

$$\dot{x}(t) = Ax(t)$$

solution: $x(t) = e^{At}x(0)$, for $t \geq 0$ (use $\dfrac{de^{At}}{dt} = Ae^{At} = e^{At}A$)

Applications to difference equations

we solve the difference equations for $t = 0, 1, \ldots$ where $x(0)$ is given

scalar: $x(t) \in \mathbf{R}$ and $a \in \mathbf{R}$ is given

$$x(t+1) = ax(t)$$

solution: $x(t) = a^t x(0)$, for $t = 0, 1, 2, \ldots$

vector: $x(t) \in \mathbf{R}^n$ and $A \in \mathbf{R}^{n \times n}$ is given

$$x(t+1) = Ax(t)$$

solution: $x(t) = A^t x(0)$, for $t = 0, 1, 2, \ldots$

example: solve the ODE

$$\ddot{y}(t) - \dot{y}(t) - 6y(t) = 0, \quad y(0) = 1, \dot{y}(0) = 0$$

solution: define

$$x(t) \triangleq \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \triangleq \begin{bmatrix} y(t) \\ \dot{y}(t) \end{bmatrix}$$

write the equation into the vector form $\dot{x}(t) = Ax(t)$

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} \dot{y}(t) \\ \ddot{y}(t) \end{bmatrix} = \begin{bmatrix} \dot{y}(t) \\ \dot{y}(t) + 6y(t) \end{bmatrix} \\ &= \begin{bmatrix} x_2(t) \\ 6x_1(t) + x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 6 & 1 \end{bmatrix} x(t) \end{aligned}$$

the initial condition is

$$x(0) = \begin{bmatrix} y(0) \\ \dot{y}(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

thus it is left to compute e^{At}

$$A = \begin{bmatrix} 0 & 1 \\ 6 & 1 \end{bmatrix}$$

the eigenvalues and eigenvectors of A are

$$\lambda_1 = -2,\ v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \lambda_2 = 3,\ v_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

all eigenvalues are distinct, so A is diagonalizable and

$$e^{At} = Te^{\Lambda t}T^{-1}, \quad T = \begin{bmatrix} v_1 & v_2 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$e^{At} = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} e^{-2t} & 0 \\ 0 & e^{3t} \end{bmatrix} \frac{1}{5} \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}$$

the closed-form expression of e^{At} is

$$e^{At} = \frac{1}{5}\begin{bmatrix} 3e^{-2t} + 2e^{3t} & -e^{-2t} + e^{3t} \\ -6e^{-2t} + 6e^{3t} & 2e^{-2t} + 3e^{3t} \end{bmatrix}$$

the solution to the vector equation is

$$\begin{aligned} x(t) = e^{At}x(0) &= \frac{1}{5}\begin{bmatrix} 3e^{-2t} + 2e^{3t} & -e^{-2t} + e^{3t} \\ -6e^{-2t} + 6e^{3t} & 2e^{-2t} + 3e^{3t} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} 3e^{-2t} + 2e^{3t} \\ -6e^{-2t} + 6e^{3t} \end{bmatrix} \end{aligned}$$

hence the solution $y(t)$ can be obtained by

$$y(t) = x_1(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) = \frac{1}{5}\left(3e^{-2t} + 2e^{3t}\right), \quad t \geq 0$$

example: solve the difference equation

$$y(t+2) - y(t+1) - 6y(t) = 0, \quad y(0) = 1, y(1) = 0$$

solution: define

$$x(t) \triangleq \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \triangleq \begin{bmatrix} y(t) \\ y(t+1) \end{bmatrix}$$

write the equation into the vector form $x(t+1) = Ax(t)$

$$\begin{aligned} x(t+1) &= \begin{bmatrix} y(t+1) \\ y(t+2) \end{bmatrix} = \begin{bmatrix} y(t+1) \\ y(t+1) + 6y(t) \end{bmatrix} \\ &= \begin{bmatrix} x_2(t) \\ 6x_1(t) + x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 6 & 1 \end{bmatrix} x(t) \end{aligned}$$

the initial condition is

$$x(0) = \begin{bmatrix} y(0) \\ y(1) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

thus it is left to compute A^t

$$A = \begin{bmatrix} 0 & 1 \\ 6 & 1 \end{bmatrix}$$

the eigenvalues and eigenvectors of A are

$$\lambda_1 = -2,\ v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \lambda_2 = 3,\ v_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

all eigenvalues are distinct, so A is diagonalizable and

$$A^t = T\Lambda^t T^{-1}, \quad T = \begin{bmatrix} v_1 & v_2 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$A^t = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} (-2)^t & 0 \\ 0 & 3^t \end{bmatrix} \frac{1}{5} \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}$$

the closed-form expression of A^t is

$$A^t = \frac{1}{5}\begin{bmatrix} 2(3^t) + 3(-2)^t & 3^t - (-2)^t \\ 2(3^{t+1}) + 3(-2)^{t+1} & 3^{t+1} - (-2)^{t+1} \end{bmatrix}$$

for $t = 0, 1, 2, \ldots$

the solution to the vector equation is

$$\begin{aligned} x(t) = A^t x(0) &= \frac{1}{5}\begin{bmatrix} 2(3^t) + 3(-2)^t & 3^t - (-2)^t \\ 2(3^{t+1}) + 3(-2)^{t+1} & 3^{t+1} - (-2)^{t+1} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} 2(3^t) + 3(-2)^t \\ 2(3^{t+1}) + 3(-2)^{t+1} \end{bmatrix} \end{aligned}$$

hence the solution $y(t)$ can be obtained by

$$y(t) = x_1(t) = \frac{1}{5}\left(2(3^t) + 3(-2)^t\right), \quad t = 0, 1, 2, \ldots$$

MATLAB commands

- expm(A) computes the matrix exponential e^A
- exp(A) computes the exponential of the entries in A

example from page 5-18, $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad e^A = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}$

```
>> A=[1 1;0 0];
>> expm(A)
ans =
    2.7183    1.7183
         0    1.0000
>> exp(A)
ans =
    2.7183    2.7183
    1.0000    1.0000
```

References

Chapter 21 in

M. Dejnakarin, Mathematics for Electrical Engineers, 3rd edition, Chulalongkorn University Press, 2006

Lecture note on

Linear algebra, EE263, S. Boyd, Stanford university