Matrix functions

Note Title 2020-03-12

$$J = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix} \in \mathbb{R}^{n \times n} \qquad J^2 = \begin{bmatrix} 0 & 0 & 1 & & \\ & 0 & 0 & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ & & & & 0 \end{bmatrix} \qquad J^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & & 0 \\ 0 & 0 & 0 & & 1 & \\ & & \ddots & & \ddots & \ddots \\ & & & & \ddots & 1 \end{bmatrix}$$

$$J^{n-1} = \begin{bmatrix} & & 1 \\ & 0 & \\ & & \end{bmatrix} \qquad J^n = 0$$

$$p(x) = p_0 + p_1 x + \dots + p_d x^d$$

$$p(J) = p_0 I + p_1 J + p_2 J^2 + \dots + p_d J^d =$$

$$= \begin{bmatrix} p_0 & p_1 & p_2 & \cdots & p_{n-1} \\ & p_0 & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & p_2 \\ & & & \ddots & p_1 \\ 0 & & & & p_0 \end{bmatrix}$$ Toeplitz (constant entries on diagonals) & upper triangular

---

$$J_\lambda = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \in \mathbb{R}^{n \times n}$$ Jordan block with eigenvalue λ

$$J_\lambda - \lambda I = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

$$p(J_\lambda) = ?$$

$$p(x) = p(\lambda) + p'(\lambda)(x-\lambda) + \frac{p''(\lambda)}{2}(x-\lambda)^2 + \frac{p'''(\lambda)}{3!}(x-\lambda)^3 + \dots + \frac{p^{(d)}(\lambda)}{d!}(x-\lambda)^d$$

$$p(J_\lambda) = p(\lambda)I + p'(\lambda)(J_\lambda - \lambda I) + \frac{p''(\lambda)}{2}(J_\lambda - \lambda I)^2 + \dots + \frac{p^{(d)}(\lambda)}{d!}(J_\lambda - \lambda I)^d$$

$$= p(\lambda)I + p'(\lambda)J_0 + \frac{p''(\lambda)}{2}J_0^2 + \dots + \frac{p^{(d)}(\lambda)}{d!}J_0^d$$

$$= \begin{bmatrix} p(\lambda) & p'(\lambda) & \frac{p''(\lambda)}{2} & \cdots & \frac{p^{(n-1)}(\lambda)}{(n-1)!} \\ & p(\lambda) & \ddots & \ddots & \frac{p''(\lambda)}{2} \\ & & \ddots & \ddots & p'(\lambda) \\ & 0 & & & p(\lambda) \end{bmatrix}$$

If $A = SJS^{-1}$ is a Jordan form of A, with $J = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_s \end{bmatrix}$

$$p(A) = p_0 I + p_1 A + \dots + p_d A^d = S\left(p_0 I + p_1 J + \dots + p_d J^d\right) S^{-1}$$

$$= S \begin{bmatrix} \boxed{\begin{matrix} p(\lambda_1) & p'(\lambda_1) & \cdots & \frac{p^{(n_1-1)}}{(n_1-1)!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & p'(\lambda_1) \\ & & & p(\lambda_1) \end{matrix}} & & & & 0 \\ & \boxed{p(J_2)} & & & \\ & & \boxed{p(J_3)} & & \\ & & & \ddots & \\ 0 & & & & \boxed{p(J_s)} \end{bmatrix} S^{-1}$$

__Idea__: Let us use the same definition for an arbitrary function

__Definition__: given $f: U \subseteq \mathbb{C} \to \mathbb{C}$, we say f is __defined on A__ if f is defined and $n_i - 1$ times differentiable on each eigenvalue λ_i of A (with maximum size of associated Jordan block n_i).

__Definition__ (attempted): Given $f: U \subseteq \mathbb{C} \to \mathbb{C}$ defined on A, and A with Jordan form $A = S \operatorname{diag}(J_1, J_2, \dots J_s) S^{-1}$,

then

$$f(A) := S \operatorname{diag}(f(J_1), \dots f(J_s)) S^{-1}, \text{ where}$$

$$f(J_i) := \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(n_i-1)!} f^{(n_i-1)}(\lambda_i) \\ & \ddots & \ddots & \vdots \\ 0 & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{bmatrix}$$

$f(x) = \sin(x)$

$f(x) = \sqrt{x}$

$f(x) = \exp(x)$

<u>Concern:</u> in the Jordan form, S is not uniquely defined. Is this definition well-posed?

Alternate definition (which is well-posed):

Given f, A s.t. f is defined on A

Take a polynomial p such that

$$\begin{cases} p(\lambda_i) = f(\lambda_i) \\ p'(\lambda_i) = f'(\lambda_i) \\ \vdots \\ p^{(n_i-1)}(\lambda_i) = f^{(n_i-1)}(\lambda_i) \end{cases}$$

(Interpolation conditions)

for each eigenvalue λ_i of A with maximum Jordan block size n_i

(Is there one such polynomial?)

Then, $f(A) := p(A) = p_0 I + p_1 A + \dots + p_d A^d$

1) This definition coincides with the previous one (we know what $p(A)$ is, and it coincides with $f(A)$)

2) This definition does not depend of p (only the values of $p(\lambda_i), \dots p^{(n_i-1)}(\lambda_i)$ appear in it).

3) Does not depend on S

Remark: if there are repeated eigenvalues, there are fewer than n conditions

Remark2: p depends on A:

$$\exp(A) = p_A(A)$$
$$\exp(B) = p_B(B)$$

Ex:

$$A = \begin{bmatrix} \boxed{\begin{matrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{matrix}} & & 0 \\ & \boxed{0} & \\ 0 & & \boxed{\begin{matrix} 4 & 1 \\ 0 & 4 \end{matrix}} \end{bmatrix}$$

$$f(x) = \sqrt{x}$$

$f(A) = \sqrt{A} = ?$

I want a polynomial with

$$\underbrace{p(0) = f(0) = 0 \quad p(4) = f(4) = 2 \quad p'(4) = f'(4) = 1/4 \quad p''(4) = f''(4) = -\tfrac{1}{32}}$$

Then,

$$f(A) = p(A) = \begin{bmatrix} p(4) & p'(4) & \frac{p''(4)}{2} & & & \\ 0 & p(4) & p'(4) & & & \\ 0 & 0 & p(4) & & & \\ & & & \boxed{p(0)} & & \\ & & & & p(4) & p'(4) \\ & & & & 0 & p(4) \end{bmatrix} = \begin{bmatrix} 2 & 1/4 & -1/64 & & & \\ & 2 & 1/4 & & & \\ & & 2 & & & \\ & & & \boxed{0} & & \\ & & & & 2 & 1/4 \\ & & & & 0 & 2 \end{bmatrix}$$

(Note that $[f(A)]^2 = A$)

Also, note that f is defined in A

(If we had $[0,1]$, we'd have a problem, because $f'(0)$ is not defined $\left(f'(x) = \frac{1}{2}\frac{1}{\sqrt{x}} \right)$

How do I compute p explicitly?

I have 4 conditions, let us look for

$p(x) = a + bx + cx^2 + dx^3 \qquad p'(x) = b + 2cx + 3dx^2$

$p''(x) = 2c + 6dx$

$p(0) = 0 \quad p(4) = 2 \quad p'(4) = \frac{1}{4} \quad p''(4) = -\frac{1}{32}$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 4 & 4^2 & 4^3 \\ 0 & 1 & 2\cdot 4 & 3\cdot 4^2 \\ 0 & 0 & 2 & 6\cdot 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1/4 \\ -1/32 \end{bmatrix} \qquad (*)$$

$$p(x) = \frac{3}{256}x^3 - \frac{5}{32}x^2 + \frac{15}{16}x.$$

$$f(A) = p(A) = \frac{3}{256}A^3 - \frac{5}{32}A^2 + \frac{15}{16}A = \begin{bmatrix} \text{same matrix as before} \end{bmatrix}$$

Hermite interpolation

Given points $x_1, x_2 \dots, x_n$, and $y_{i,0}, \dots y_{i,m_i-1}$, $i=1, \dots n$, multiplicities $m_1, m_2, \dots m_n$, there exists a unique polynomial of degree

$d < m_1 + m_2 + \dots + m_n =: m$ such that

$$p(x_i) = y_{i,0}, \quad p'(x_i) = y_{i,1}, \quad \dots \quad p^{(m_i-1)}(x_i) = y_{i,m_i-1} \quad i = 1, 2 \dots n$$

Proof: This is a linear system in the coefficients of the polynomial, like in (*).

I just have to prove that the associated "generalized Vandermonde" matrix is invertible:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\ & 1 & 2x_1 & \cdots & (m-1)x_1^{m-2} \\ & & \ddots & & \vdots \\ \hline 1 & x_2 & x_2^2 & \cdots & \\ & & & \ddots & \\ & & \vdots & & \end{bmatrix} \qquad z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} \qquad Vz = 0$$

Let us take $z \in \ker V$ $\quad Vz = 0$ Define $z(x) = z_1 + z_2 x + \cdots + z_m x^{m-1}$

Then, $Vz = 0 \iff$

$$\begin{array}{cccc} z(x_1) = 0 & z(x_2) = 0 & & z(x_n) = 0 \\ z'(x_1) = 0 & \vdots & \cdots & \vdots \\ \vdots & & & \\ z^{(m_1-1)}(x_1) = 0 & z^{(m_2-1)}(x_1) = 0 & & z^{(m_n-1)}(x_n) = 0 \end{array}$$

$$\overset{=0}{z}(x) = (x-x_1)^{m_1}(x-x_2)^{m_2} \cdots (x-x_n)^{m_n} \cdot q(x)$$

But $\deg z(x) < m_1 + \dots + m_n$, so $z(x) \equiv 0 \Rightarrow \operatorname{Ker} V = \{0\}$.

Matlab:

sqrt(A)	elementwise	exp(A)
sqrtm(A)	matrix function	expm(A)

<u>Example</u>:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad f(x) = \sqrt{x}$$

f is <u>not</u> defined in A, because $f'(0)$ does not exist.

Indeed, there is no matrix X s.t. $X^2 = A$:

if there were one, it would have to have two zero eigenvalues so Jordan form $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, but in both cases $x^2 = 0$.

$$A = S \begin{bmatrix} \boxed{-1} & & \\ & \boxed{0} & \\ & & \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}} \end{bmatrix} S^{-1}$$

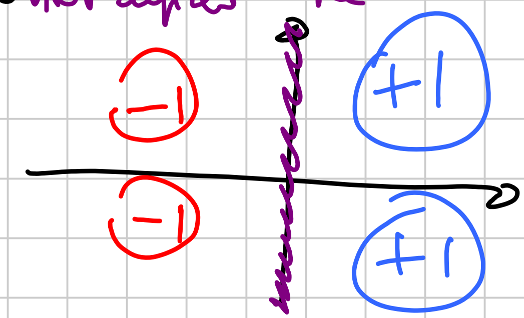
$$f(x) = \exp(x)$$

$$\exp(A) = 1 + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \dots$$

we have not proved that it coincides with the Taylor series

$$\exp(A) = S \begin{bmatrix} \boxed{e^{-1}} & & \\ & \boxed{1} & \\ & & \boxed{\begin{matrix} e & \underbrace{\boxed{e}}_{f'(1)} \\ 0 & e \end{matrix}} \end{bmatrix} S^{-1}$$

$$f(x) = \operatorname{sign}(x) = \begin{cases} 1 & \operatorname{Re} x > 0 \\ -1 & \operatorname{Re} x < 0 \\ \text{not defined} & \operatorname{Re} x = 0 \end{cases}$$



$$A = S \begin{bmatrix} \boxed{-3} & & & \\ & \boxed{-2} & & \\ & & 1 & 1 \\ & & 0 & 1 \end{bmatrix} S^{-1}$$

$$\operatorname{sign}(A) = S \begin{bmatrix} \boxed{-1} & & & \\ & \boxed{-1} & & \\ & & 1 & 0 \\ & & 0 & 1 \end{bmatrix} S^{-1}$$

$f'(1)$

$$\ker\left(\operatorname{sign}(A) - I\right) = \ker\left(S\left(\begin{bmatrix} -1 & & & \\ & -1 & & \\ & & 1 & 0 \\ & & 0 & 1 \end{bmatrix} - I\right)S^{-1}\right) =$$

$$\ker\left(S\begin{bmatrix} -2 & & & \\ & -2 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}S^{-1}\right) = \operatorname{span}(S_3, S_4)$$

last two columns of S

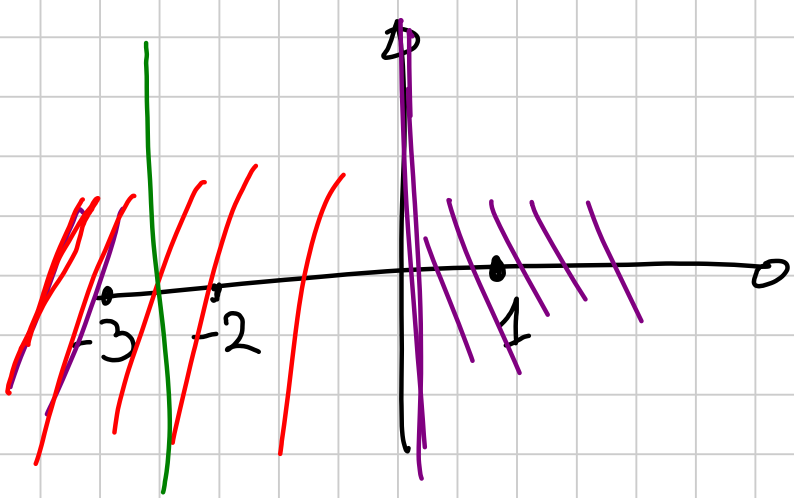
$$\ker\left(\underset{\sim}{\operatorname{sign}}(A) + I\right) = \ker\left(S\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 2 & \\ & & & 2 \end{bmatrix}S^{-1}\right) = \operatorname{span}(S_1, S_2)$$

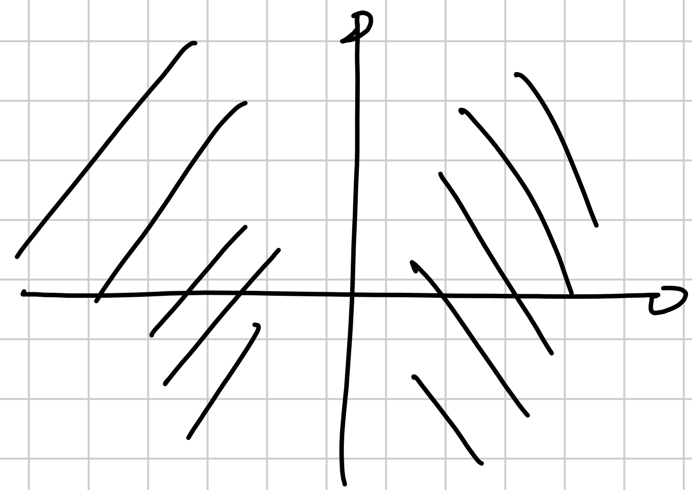
first two columns of S

$\ker(\operatorname{sign}(A) - I) = \operatorname{span}(\text{Jordan chains associated to eigenvalues with } \mathrm{Re}(\lambda) > 0)$

$\ker(\operatorname{sign}(A) + I) = \operatorname{span}(\text{Jordan chains } \mathrm{Re}(\lambda) < 0)$

(two invariant subspaces of A)



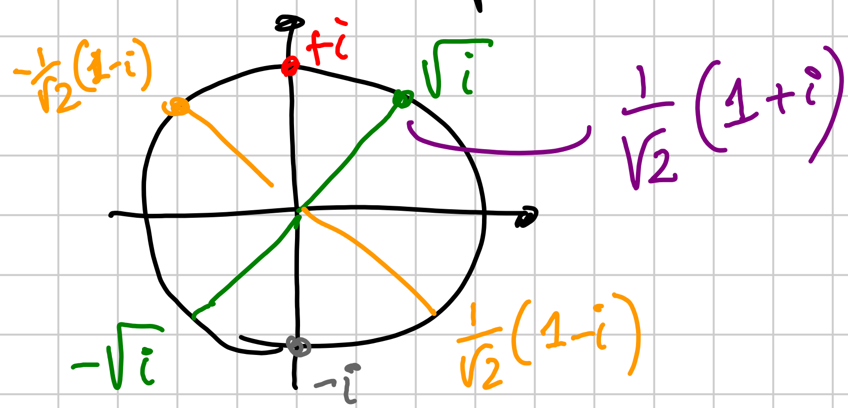


example

$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ $\quad f(x) = \sqrt{x}$ $\quad$ what is $f(A)$?

It depends?

$eig(A) = \{+i, -i\}$



We need to choose branches

for each eigenvalue to have $f(x)$ fully defined

Usual choice: the root in the right half-plane (principal square root)

$f(i) = \frac{1}{\sqrt{2}}(1+i)$

$f(-i) = \frac{1}{\sqrt{2}}(1-i)$

$\rightarrow$ interp. polynomial: $p(i) = \frac{1}{\sqrt{2}}(1+i)$, $p(-i) = \frac{1}{\sqrt{2}}(1-i)$ $\qquad p(x) = \frac{1}{\sqrt{2}}(1+x)$

$f(A) = p(A) = \frac{1}{\sqrt{2}}(I+A) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ principal square root of A

Remark: with non-conjugated choices for $\tilde{f}(i)$, $\tilde{f}(-i)$

$\tilde{f}(i) = \frac{1}{\sqrt{2}}(1+i)$

$\tilde{f}(-i) = -\frac{1}{\sqrt{2}}(1-i)$

then $f(A)$ is not real

$p(x) = \frac{1}{\sqrt{2}}(i - ix)$

$\tilde{f}(A) = \frac{1}{\sqrt{2}} i(I-A) = \frac{1}{\sqrt{2}} i \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

$[\tilde{f}(A)]^2 = A$

$[f(A)]^2 = A$

$$A = S \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix} S^{-1} \qquad f(x) = \sqrt{x}$$

Can I choose branches to obtain

NOT A MATRIX FUNCTION
(WITH OUR DEFINITION)

$$f(A) = S \begin{bmatrix} 1 & & \\ & -1 & \\ & & \sqrt{2} \end{bmatrix} S^{-1} \qquad [f(A)]^2 = S \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix} S^{-1} = A$$

No! With our definition, $f(\lambda_i)$ has to be defined uniquely on all distinct λ_i's "$f(1) = f(1)$" (In interp. polynomials, I have to choose $p(1)$ uniquely)

These are sometimes called <u>non-primary matrix functions</u>

Given a multi-valued function $f(x)$ (something with branches: $\sqrt{x}$, $\log(x)$, ...)

a nonprimary function of $A = SJS^{-1}$ is obtained by taking

$$f(A) = S \begin{bmatrix} f(J_1) & & & \\ & f(J_2) & & \\ & & \ddots & \\ & & & f(J_s) \end{bmatrix} S^{-1}$$

, where each $f(J_i)$ is constructed taking possibly a different branch of f.

This <u>does</u> depend on the choice of S.

<u>Example</u>: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $f(x) = \sqrt{x}$ Jordan form: $A = S \cdot \begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \end{bmatrix} \cdot S^{-1}$

with any invertible S. All matrices of the form

$B = S\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}S^{-1}$, for all S, are nonprimary square roots of A. They all satisfy $B^2 = I$

They are <u>not</u>, in general, polynomials in A (because $p(A)$ is a scalar multiple of I for each p)