# Matrix Functions

Note Title 2025-03-13

$$f : U \subset \mathbb{C} \to \mathbb{C} \qquad A \in \mathbb{C}^{n \times n}$$

$$A = V \operatorname{diag}\left(J_{\lambda_1}, J_{\lambda_2}, \dots J_{\lambda_s}\right) V^{-1}$$

$$f(A) = V \operatorname{diag}\left(f(J_{\lambda_1}), f(J_{\lambda_2}), \dots f(J_{\lambda_s})\right) V^{-1}$$

$$f(J_{\lambda_i}) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2} f''(\lambda_i) & \cdots & \frac{1}{(k_i-1)!} f^{(k_i-1)}(\lambda_i) \\ & \ddots & \ddots & \ddots & \vdots \\ & & & & \\ 0 & & & & f(\lambda_i) \end{bmatrix}$$

**Or**: given a polynomial $p$ such that

$$f(\lambda_i) = p(\lambda_i) \quad f'(\lambda_i) = p'(\lambda_i) \quad \dots \quad f^{(k_i-1)}(\lambda_i) = p^{(k_i-1)}(\lambda_i)$$

$\forall i = 1, \dots s$

then $f(A) = p(A)$

**Example**:

$$A = \begin{bmatrix} 4 & 1 & 0 & \\ 0 & 4 & 1 & \\ 0 & 0 & 4 & 1 \\ & & & 10 \end{bmatrix} \qquad f(x) = \sqrt{x}$$

$$f(4) = 2 \qquad f'(x) = \frac{1}{2\sqrt{x}} \qquad f'(4) = \frac{1}{4} \qquad f''(x) = -\frac{1}{4} x^{-3/2}$$

$$f''(4) = -\frac{1}{4 \cdot 8} = -\frac{1}{32}$$

Version 1:

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

Version 2: let us find $p(x)$ s.t.

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

We try looking for

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

It is a linear system: $\qquad p''(x)=2c+6xd$

$$\underbrace{\begin{bmatrix} 1 & 4 & 4^2 & 4^3 \\ 0 & 1 & 2\cdot 4 & 3\cdot 4^2 \\ 0 & 0 & 2 & 6\cdot 4 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{V} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{1}{4} \\ -\frac{1}{32} \\ 0 \end{bmatrix}$$

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

(Matlab verification of $p(A)=f(A)$)

Theorem (Hermite interpolation)

Given $\lambda_1, \lambda_2, \dots \lambda_s \in \mathbb{C}$ and multiplicities

$k_1, k_2 \dots \quad k_s \in \mathbb{N}$

If the $\lambda_i$ are distinct, there exists a unique polynomial of degree $d < k_1+k_2+\dots+k_s$ such that

$$p(\lambda_i)=y_{i,0} \quad p'(\lambda_i)=y_{i1}, \dots \quad p^{(k_i-1)}(\lambda_i)=y_{i,k_i-1}$$

$\forall i=1,2,\dots s$

for any choice of the $y_{ij}$.

Proof: These conditions are equivalent to a linear system $\qquad n=k_1+k_2+\dots+k_s$

$$\begin{bmatrix} & & \\ & V & \\ & & \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{n-1} \end{bmatrix} = \begin{bmatrix} \\ y \\ \\ \end{bmatrix}$$

We would like to prove that $V$ is invertible.
Suppose there is a vector $q$ such that

$$\begin{bmatrix} & & \\ & V & \\ & & \end{bmatrix} \cdot \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_{n-1} \end{bmatrix} = 0 \qquad \text{i.e., } q \in \ker V$$

These conditions imply that the polynomial

$$q(x) = q_0 + q_1 x + q_2 x^2 + \dots + q_{n-1} x^{n-1}$$

is such that

$$q(\lambda_i) = 0 \quad q'(\lambda_i) = 0, \quad \dots \quad q^{(k_i - 1)}(\lambda_i) = 0 \qquad \forall i = 1, 2, \dots s.$$

Hence $q(x)$ has roots in $\lambda_i$ of multiplicity $k_i$, and it must be a multiple of

$$(x - \lambda_1)^{k_1} (x - \lambda_2)^{k_2} \dots (x - \lambda_s)^{k_s}$$

which has degree $k_1 + k_2 + \dots + k_s = n$, but for degree reasons it must be the zero polynomial.

---

"Counterexample": if $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\quad f(x) = \sqrt{x}$

$f'(0)$ does not exist $\to$ this function is <u>not</u> defined in the matrix $A$.

This makes sense, because the matrix equation $X^2 = A$ has no solution $X \in \mathbb{C}^{2\times 2}$ for $A$ as above.

Such a solution $X$ must have only $0$ as an eigenvalue: indeed, if $Xv = v\lambda$, then $X^2 v = v\lambda^2$ $\quad X^2 = A$
$\Rightarrow \lambda^2$ must be an eigenvalue of $A$ $\Rightarrow \lambda = 0$

So $X$ must have Jordan form either

$$X = V \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} V^{-1} \qquad \text{or} \qquad X = V \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} V^{-1} = 0$$

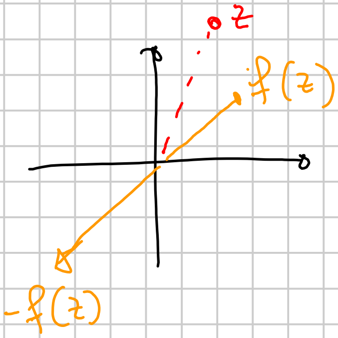
$X^2 = V\begin{bmatrix}0&1\\0&0\end{bmatrix}^2 V^{-1} = 0$ $\qquad$ $X=0$, impossible

Neither option produces a valid $X$.

Example $\qquad f(x)=\sqrt{x} \qquad A=\begin{bmatrix}0&1\\-1&0\end{bmatrix} \qquad \Lambda(A)=\{i,-i\}$

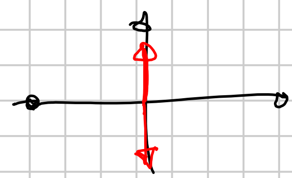
We must choose a branch of the square root to define $f(x)$ properly.

Most common choice: principal square root: the one such that $f(x) \in$ right half plane

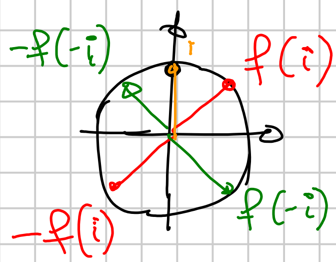



For any $z$ that is not a negative real, there exists exactly one square root of $z$ in the right half-plane $\{z : \operatorname{Re} z > 0\}$

(Let us leave it undefined on the negative reals)



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




An interpolating polynomial is $p(x) = \frac{\sqrt{2}}{2}(1+x)$

$$f(A) = p(A) = \frac{\sqrt{2}}{2}\left(\begin{bmatrix}1&0\\0&1\end{bmatrix} + \begin{bmatrix}0&1\\-1&0\end{bmatrix}\right) = \frac{\sqrt{2}}{2}\begin{bmatrix}1&1\\-1&1\end{bmatrix}.$$

---

We can choose a different branch, obtaining a function $g(x)$

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

$g(x)$ has interpolating polynomial

$p(x) = \frac{\sqrt{2}}{2} i(1-x)$, the unique polynomial s.t. $\begin{cases} p(i) = g(i) \\ p(-i) = g(-i) \end{cases}$

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

is another matrix function associated to a different branch of the square root.

---

$f(x) = e^x$

$$A = V\begin{bmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & \\ & & & -1 \end{bmatrix}V^{-1}$$

$$f(A) = V\begin{bmatrix} \boxed{\begin{matrix} e^2 & e^2 \\ 0 & e^2 \end{matrix}} & & \\ & \boxed{e^2} & \\ & & \boxed{e^{-1}} \end{bmatrix}V^{-1}$$

Alternatively, look for $p$ s.t.

$\boxed{\begin{matrix} p(2) = e^2 \\ p(2) = e^2 \end{matrix}}$ $\quad p'(2) = e^2$

$p(-1) = e^{-1}$

Any polynomial s.t. $p(2) = e^2$, $p'(2) = e^2$, $p(-1) = e^{-1}$ is such that

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

We can find a polynomial of degree 2, in this case, but nothing changes.

If we want to compute with a function that has multiple branches, such as the square root, we need to choose a branch first:

$$f(2) = -\sqrt{2}$$

$$f(-1) = i$$

then we can apply the definitions. In particular, with our definition, all Jordan blocks have the same branch:

$$f(A) = V \begin{bmatrix} -\sqrt{2} & \frac{1}{2\sqrt{2}} & & \\ 0 & -\sqrt{2} & & \\ & & -\sqrt{2} & \\ & & & i \end{bmatrix} V^{-1}$$

This is part of our definition, $f(A_i)$ must be defined uniquely!

The matrix

$$B = V \begin{bmatrix} -\sqrt{2} & \frac{1}{2\sqrt{2}} & & \\ & -\sqrt{2} & & \\ & & +\sqrt{2} & \\ & & & i \end{bmatrix} V^{-1} \qquad A = \begin{bmatrix} 2 & \frac{1}{2} & & \\ & 2 & & \\ & & 2 & \\ & & & -1 \end{bmatrix}$$

is not a matrix function in our definition.

Note: there is no polynomial $p$ s.t. $p(A) = B$.

These "pseudo-functions" such as $B$ sometimes are called nonprimary functions, so $B$ would be a non-primary square root of $A$.

(Note that $B^2 = A$)

⚠ The interpolating polynomial $p$ depends on the matrix $A$ so it is not correct to claim that "all matrix functions are polynomials".

However, we can prove many properties of matrix functions by changing $f$ into $p$

- For any $f$ and $A$, $\quad Af(A) = f(A)A$

  Proof: take $p$ s.t. $p(A) = f(A)$.

- For any $f, A$, invertible $M$:

$$f(MAM^{-1}) = M f(A) M^{-1}$$

Proof: take $p$ that interpolates $f$ on $\Lambda(A)$ with the right multiplicities, $f(A) = p(A)$

but also $f(MAM^{-1}) = p(MAM^{-1})$

- $$f\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) = \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix}$$

Proof: take $p$ that interpolates $f$ on $\Lambda\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right)$

then $f\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) = p\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right)$

but also $f(A) = p(A)$, $f(B) = p(B)$

- If $h(x) = f(x) g(x)$, then $h(A) = f(A) g(A)$

$f(A) = p_f(A)$ $\quad$ $g(A) = p_g(A)$

One can verify that $p_h = p_f(x) p_g(x)$ is a polynomial that interpolates $h(x) = f(x) g(x)$.

So $\quad h(A) = p_h(A) = p_f(A) p_g(A) = f(A) g(A)$.

Similarly, $h(x) = f(x) + g(x) \Rightarrow h(A) = f(A) + g(A)$

$h(x) = f(g(x)) \Rightarrow h(A) = f(g(A))$

These properties allow us to prove that matrix functions satisfy certain identities that scalar functions do.

$f(x) = \sqrt{x}$ satisfies $f(x)^2 = x$

hence $B = f(A)$ satisfies $B^2 = A$, which we observed earlier numerically.

This holds for many other identities, e.g.

$$[\sin(A)]^2 + [\cos(A)]^2 = I.$$

More properties of matrix functions:

- If $\Lambda(A) = \{\lambda_1, \lambda_2, \dots \lambda_n\}$

$$\Lambda(f(A)) = \{f(\lambda_1), f(\lambda_2), \dots f(\lambda_n)\}$$

Algebraic multiplicities are preserved, but geometric multiplicities can change: indeed,

$$f\left(\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{bmatrix}\right) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''}{2!}(\lambda) & \frac{f'''}{3!}(\lambda) \\ & \ddots & \ddots & \ddots \\ 0 & & & \end{bmatrix}$$

If $f'(\lambda) = 0$, then

$$f(J) = \begin{bmatrix} f(\lambda) & 0 & * & * \\ 0 & f(\lambda) & 0 & * \\ 0 & 0 & f(\lambda) & 0 \\ 0 & 0 & 0 & f(\lambda) \end{bmatrix}$$

| | |
|---|---|
| $\mathrm{rk}(J-\lambda I) = 3,$ | $\mathrm{rk}\big(f(J) - f(\lambda)I\big) = 2,$ |
| $m_g(\lambda) = 1.$ | $m_g(f(\lambda)) = 2.$ |