Lecture 4: Functions of Square Matrices

Baboo J. Cui, Yangang Cao

July 9, 2019

Contents

1 Functions of Scalars

Examples of functions of a scalar $\lambda \in \mathbb{R}$:

$$f(\lambda) = \lambda^k, \sqrt{\lambda}, e^\lambda, \ln \lambda, \frac{1}{1-\lambda}$$

Can we define functions of square matrices $A \in \mathbb{R}^{n \times n}$?

$$f(A) = A^k, A^{\frac{1}{2}}, e^A, \ln(A), (l_n - A)^{-1}$$

(Real) analytic functions:

- $f(\lambda)$ is infinitely differentiable
- At any λ_0 where $f(\cdot)$ is defined, its Taylor expansion converges locally:

$$f(\lambda) = f(\lambda_0) + f'(\lambda_0)(\lambda - \lambda_0) + \frac{1}{2!} f''(\lambda_0)(\lambda - \lambda_0)^2 + \cdots$$

2 Polynomials of Square Matrices

For polynomials $f(\lambda) = a_k \lambda^k + a_{k-1} \lambda^{k-1} + \cdots + a_0$, the extension is easy:

$$f(A) = a_k A^k + a_{k-1} A^{k-1} + \cdots + a_0 l_n \in \mathbb{R}^{n \times n}, \quad \forall A \in \mathbb{R}^{n \times n}$$

- Replace every occurrence of λ by A
- Replace constant term a_0 by $a_0 I_n$

3 Polynomials of Matrices via JCF

Suppose $A \in \mathbb{R}^{n \times n}$ has Jordan Canonical Form (JCF):

$$A = TJT^{-1} = T \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{bmatrix} T^{-1}$$

For a polynomial function $f(\lambda), f(A)$ can be computed via (JCF).

4 Ploynomials of Jordan Blocks

Each Jordan block J_i of size n_i can be decomposed as $J_i = \lambda_i I + N_i$

- $\lambda_i I$ and N_i commute
- N_i is nilpotent: $(N_i)^{n_i} = 0$

Polynomial function of J_i can be computed using these properties.

5 Characteristic Polynomial of Jordan Blocks

Let $\chi_A(\lambda) = \det(\lambda l - A)$ be the characteristic polynomial of A

Lemma:
For any Jordan block J_i of $A, \chi_A(J_i) = 0$.

6 Cayley-Hamilton Theorem

Theorem (Cayley-Hamilton Theorem):
For any matrix $A \in \mathbb{R}^{n \times n}$, $\chi_A(A) = 0$.

Proof:

$$\chi_A(A) = T \begin{bmatrix} \chi_A(J_1) & & \\ & \ddots & \\ & & \chi_A(J_q) \end{bmatrix} T^{-1} = 0$$

7 Minimal Polynomials

Definition:
The minimal polynomial of $A \in \mathbb{R}^{n \times n}$ is the polynomial $\mu_A(\lambda)$ with the minimum degree satisfying $\mu_A(A) = 0$.

Using the JCF, the minimal polynomial is

$$\mu_A(\lambda) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_\ell)^{d_\ell}$$

- $\lambda_1, \ldots, \lambda_\ell$ are the distinct eigenvalues of A
- d_i is the largest size of Jordan blocks associated with λ_i

8 Implication of C-H Theorem: I

Corollary:
Given a square matrix $A \in \mathbb{R}^{n \times n}$

- $A^k \in \operatorname{span}\left\{I_n, A, A^2, \ldots, A^{n-1}\right\}$, for $k = 0, 1, 2, \ldots$
- For any polynomial function $f(\lambda)$

$$f(A) \in \operatorname{span}\left\{I_n, A, A^2, \ldots, A^{n-1}\right\}$$

$$\Rightarrow f(A) = h(A) \text{ for some polynomial } h(\lambda) \text{ of degree } \leq n-1$$

9 Implication of C-H Theorem: II

Theorem:
Given a square matrix $A \in \mathbb{R}^{n \times n}$ and a polynomial function $f(\lambda)$,

$$f(A) = h(A)$$

for a polynomial $h(\lambda)$ of degree at most $n-1$ satisfying

$$\begin{cases} f^{(j)}(\lambda_1) = h^{(j)}(\lambda_1), & j = 0, 1, \ldots, m_1 - 1 \\ f^{(j)}(\lambda_\ell) = h^{(j)}(\lambda_\ell), & j = 0, 1, \ldots, m_\ell - 1 \end{cases} \tag{1}$$

- $\lambda_1, \ldots, \lambda_\ell$ are the distinct eigenvalues of A
- $m_1, \ldots, m_\ell$ are their algebraic multiplicities
- $f(\lambda)$ and $h(\lambda)$ agree on the spectrum $\sigma(A)$ of matrix A

10 General Functions of Square Matrices

Given an analytical function $f(\lambda)$ that is defined at $\lambda = 0$

$$\begin{aligned} & f(\lambda) = f(0) + f'(0)\lambda + \frac{1}{2!} f''(0)\lambda^2 + \cdots \\ \Rightarrow \quad & f(A) = f(0)I + f'(0)A + \frac{1}{2!} f''(0)A^2 + \cdots \end{aligned}$$

Definition (Function of Square Matrix):
The general function of the square matrix $A \in \mathbb{R}^{n \times n}$ is defined as

$$f(A) = h(A)$$

where $h(\lambda)$ is a polynomial of degree at most $n-1$ that agrees with $f(\lambda)$ on the spectrum $\sigma(A)$ of A (see eq (1) for definition).

11 Example

Find $f(A) = e^A$ for $A_1 = \begin{bmatrix} 0.4 & 0.6 \\ 0.7 & 0.3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{bmatrix}$

12 General Functions of Matrices via JCF

Suppose $A \in \mathbb{R}^{n \times n}$ has Jordan Canonical Form (JCF):

$$A = T \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{bmatrix} T^{-1} \Rightarrow f(A) = T \begin{bmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_q) \end{bmatrix} T^{-1}$$

For each Jordan block J_i:

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

13 Example

Consider $A = \begin{bmatrix} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{bmatrix}$

- $f_1(\lambda) = \lambda^{1000}$, $f_1(A) = A^{1000}$
- $f_2(\lambda) = \frac{1}{s-\lambda}$ for some constant scalar s, $f_2(A) = (sI - A)^{-1}$