Polynomials of matrices

Another different way to make polynomials and matrices interact: take a scalar polynomial, and apply a (square) matrix to it, e.g.,

$$p(x) = 1 + 3x - 5x^2 \implies p(A) = I + 3A - 5A^2.$$

Lemma

If $A = S\,\mathrm{blkdiag}(J_1, J_2, \ldots, J_s)S^{-1}$ is a Jordan form, then $p(A) = S\,\mathrm{blkdiag}(p(J_1), p(J_2), \ldots, p(J_s))S^{-1}$, and

$$p(J_i) = \begin{bmatrix} p(\lambda_i) & p'(\lambda_i) & \ldots & \frac{1}{k}p^{(k)}(\lambda_i) \\ & p(\lambda_i) & \ddots & \vdots \\ & & \ddots & p'(\lambda_i) \\ & & & p(\lambda_i) \end{bmatrix}.$$

(Proof: Taylor expansion of p around λ.)

Functions of matrices [Higham book, '08]

We can extend the same definition to arbitrary scalar functions:

Given a function $f : U \subseteq \mathbb{C} \to \mathbb{C}$, we say that f is defined on A if f is defined and differentiable at least $m_g(\lambda_i) - 1$ times on each eigenvalue λ_i of A.

Definition

If $A = S \operatorname{blkdiag}(J_1, J_2, \dots, J_s) S^{-1}$ is a Jordan form, then $f(A) = S \operatorname{blkdiag}(f(J_1), f(J_2), \dots, f(J_s)) S^{-1}$, where

$$f(J_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \dots & \frac{1}{k} f^{(k)}(\lambda_i) \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{bmatrix}.$$

(Reasonable doubt: is it independent of the choice of S?)

Alternate definition: via Hermite interpolation

Definition

$f(A) = p(A)$, where p is a polynomial such that $f(\lambda_i) = p(\lambda_i), f'(\lambda_i) = p'(\lambda_i), \dots, f^{(m_g(\lambda_i)-1)}(\lambda_i) = p^{(m_g(\lambda_i)-1)}(\lambda_i)$ for each i.

We may use this as a definition of $f(A)$ (and it does not depend on S).

Obvious from the definitions that it coincides with the previous one.

Remark: if $A \in \mathbb{C}^{m \times m}$ has multiple Jordan blocks with the same eigenvalue, these may be fewer than m conditions.

Remark: be careful when you say "all matrix functions are polynomials", because p depends on A.

Some properties

- If the eigenvalues of A are $\lambda_1, \dots, \lambda_s$, the eigenvalues of $f(A)$ are $f(\lambda_1), \dots, f(\lambda_s)$. (Remark: geometric multiplicities may drop)
- $f(A)g(A) = g(A)f(A) = (fg)(A)$ (since they are all polynomials in A).
- If $f_n \to f$ together with 'enough derivatives' (for instance because they are analytic and the convergence is uniform), then $f_n(A) \to f(A)$.
- If a sequence of matrix $A_n \to A$, then $f(A_n) \to f(A)$.
  Proof: let p_n be the (Hermite) interpolating polynomial on the eigenvalues of A_n. Interpolating polynomials are continuous in the nodes, so $p_n \to p$ (coefficient by coefficient). Then $\|p_n(A_n) - p(A)\| \leq \|p_n(A_n) - p_n(A)\| + \|p_n(A) - p(A)\| \leq \dots$.

Examples

- $\exp(A) = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \dots$. The identity holds also for non-diagonalizable matrices (by continuity).
- $\operatorname{sgn}(A)$ from

$$\operatorname{sgn}(x) = \begin{cases} -1 & \operatorname{Re} x < 0, \\ 1 & \operatorname{Re} x > 0, \\ \text{undefined} & \operatorname{Re} x = 0. \end{cases}$$

- $A^{1/2}$ from $f(x) = \sqrt{x}$. Note that we can choose signs (branch) independently on each eigenvalue. All the various ways satisfy $(A^{1/2})^2 = A$.

Nonprimary matrix functions

If a matrix A has multiple eigenvalues, one could also define a 'square root' by choosing different signs on Jordan blocks with the same eigenvalue: for instance, $\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$ as a square root of I_2 (or also $V \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} V^{-1}$ for any invertible $V \dots$).

These are called nonprimary matrix functions (and they are not matrix functions with our definition).

(They all satisfy $f(A)^2 = A$.)

(They are not polynomials in A).

Cauchy integrals

If f is analytic on and inside a contour Γ that encloses the eigenvalues of A,

$$f(A) = \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1} dz.$$

Generalizes the analogous scalar formula.

Proof If $A = V\Lambda V^{-1} \in \mathbb{C}^{m\times m}$ is diagonalizable, the integral equals

$$V \begin{bmatrix} \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-\lambda_1}dz & & \\ & \ddots & \\ & & \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-\lambda_m}dz \end{bmatrix} V^{-1} = V \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_m) \end{bmatrix} V^{-1}.$$

By continuity, the equality holds also for non-diagonalizable A.

Methods

Matrix functions arise in several areas: solving ODEs (e.g. $\exp A$), matrix analysis (square roots), physics, ...

Main methods to compute them:

- Factorizations (eigendecompositions, Schur...),
- Matrix versions of scalar iterations (e.g., Newton on $x^2 = a$),
- Interpolation / approximation,
- Complex integrals.