Lecture I: Review of Matrix Theory and Matrix Functions

Michele Benzi
Department of Mathematics and Computer Science
Emory University
Atlanta, Georgia, USA

Summer School on Theory and Computation of Matrix Functions
Dobbiaco, 15-20 June, 2014
Outline

1. Plan of these lectures
2. Review of matrix theory
3. Functions of matrices
4. Bibliography
Outline

1. Plan of these lectures
2. Review of matrix theory
3. Functions of matrices
4. Bibliography
Plan and acknowledgements

- Lecture I: Review of Matrix Theory and Matrix Functions
- Lecture II: Matrix Functions in Network Science, Part 1
- Lecture III: Matrix Functions in Network Science, Part 2
- Lecture IV: Decay in Functions of Sparse Matrices
- Lecture V: Matrix Functions in Quantum Chemistry
- Lecture VI: Numerical Analysis of Quantum Graphs

Many thanks to the Organizers of the School and to my collaborators Paola Boito, Ernesto Estrada, Christine Klymko, Nader Razouk and Mario Arioli. I would also like to gratefully acknowledge the early influence of Gene Golub.
Gene Howard Golub (1932–2007)

Ghent, Belgium, September 2006 (courtesy of Gérard Meurant)
1. Plan of these lectures

2. Review of matrix theory

3. Functions of matrices

4. Bibliography
Matrix classes

Unless otherwise specified, only matrices with entries in $\mathbb{R}$ or $\mathbb{C}$ will be considered.

A matrix $A \in \mathbb{C}^{n \times n}$ is

- **Hermitian** if $A^* = A$
- **skew-Hermitian** if $A^* = -A$
- **unitary** if $A^* = A^{-1}$
- **symmetric** if $A^T = A$
- **skew-symmetric** if $A^T = -A$
- **orthogonal** if $A^T = A^{-1}$
- **normal** if $AA^* = A^*A$

**Theorem:** $A \in \mathbb{C}^{n \times n}$ is normal if and only if there exist $U \in \mathbb{C}^{n \times n}$ unitary and $D \in \mathbb{C}^{n \times n}$ diagonal such that $U^* AU = D$. 
Jordan Normal Form

Any matrix \( A \in \mathbb{C}^{n \times n} \) can be reduced to the form

\[
Z^{-1}AZ = J = \text{diag} \left( J_1, J_2, \ldots, J_p \right),
\]

where \( Z \in \mathbb{C}^{n \times n} \) is nonsingular and \( n_1 + n_2 + \ldots + n_p = n \). The Jordan matrix \( J \) is unique up to the ordering of the blocks, but \( Z \) is not. The \( \lambda_k \)'s are the eigenvalues of \( A \). These constitute the spectrum of \( A \), denoted by \( \sigma(A) \).

**Definition:** The order \( n_i \) of the largest Jordan block in which the eigenvalue \( \lambda_i \) appears is called the index of \( \lambda_i \).
Diagonalizable matrices

A matrix $A \in \mathbb{C}^{n \times n}$ is diagonalizable if there exists an invertible matrix $X \in \mathbb{C}^{n \times n}$ such that $X^{-1}AX = D$ is diagonal. In this case all the Jordan blocks are $1 \times 1$.

From $AX = XD$ it follows that the columns of $X$ are corresponding eigenvectors of $A$, which form a basis for $\mathbb{C}^n$.

Normal matrices are precisely those matrices that can be diagonalized by unitary transformations. Thus: a matrix $A$ is normal if and only if there exists an orthonormal basis for $\mathbb{C}^n$ consisting of eigenvectors of $A$.

The eigenvalues of a normal matrix can lie anywhere in $\mathbb{C}$. Hermitian matrices have real eigenvalues; skew-Hermitian matrices have purely imaginary eigenvalues; and unitary matrices have eigenvalues of unit modulus, i.e., if $\lambda \in \sigma(U)$ with $U$ unitary then $|\lambda| = 1$. 
Some useful expressions

From the Jordan decomposition of a matrix $A \in \mathbb{C}^{n \times n}$ we obtain the following “coordinate-free” decomposition of $A$:

$$A = \sum_{i=1}^{s} \left[ \lambda_i G_i + N_i \right]$$

where $\lambda_1, \ldots, \lambda_s$ are the distinct eigenvalues of $A$, $G_i$ is the projector onto the generalized eigenspace $\text{Ker}((A - \lambda_i I)^{n_i})$ along $\text{Ran}((A - \lambda_i I)^{n_i})$ with $n_i = \text{index}(\lambda_i)$, and $N_i = (A - \lambda_i I)G_i = G_i(A - \lambda_i I)$ is nilpotent of index $n_i$. The $G_i$’s are the Frobenius covariants of $A$.

If $A$ is diagonalizable ($A = XDX^{-1}$) then $N_i = 0$ and the expression above can be written

$$A = \sum_{i=1}^{n} \lambda_i x_i y_i^*$$

where $\lambda_1, \ldots, \lambda_n$ are not necessarily distinct eigenvalues, and $x_i, y_i$ are right and left eigenvectors of $A$ corresponding to $\lambda_i$. Hence, $A$ is a weighted sum of at most $n$ rank-one matrices (oblique projectors).
If \( A \) is normal then the spectral theorem yields

\[
A = \sum_{i=1}^{n} \lambda_i u_i u_i^* \]

where \( u_i \) is eigenvector corresponding to \( \lambda_i \). Hence, \( A \) is a weighted sum of at most \( n \) rank-one orthogonal projectors.

From these expressions one readily obtains for any matrix \( A \in \mathbb{C}^{n \times n} \) that

\[
\text{Tr}(A) := \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i
\]

and, more generally,

\[
\text{Tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k, \quad \forall k = 1, 2, \ldots
\]
Singular Value Decomposition (SVD)

For any matrix $A \in \mathbb{C}^{m \times n}$ there exist unitary matrices $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ and a “diagonal” matrix $\Sigma \in \mathbb{R}^{m \times n}$ such that

$$U^* AV = \Sigma = \text{diag} (\sigma_1, \ldots, \sigma_p)$$

where $p = \min\{m, n\}$.

The $\sigma_i$ are the singular values of $A$ and satisfy

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0,$$

where $r = \text{rank}(A)$. The matrix $\Sigma$ is uniquely determined by $A$, but $U$ and $V$ are not.

The columns $u_i$ and $v_i$ of $U$ and $V$ are left and right singular vectors of $A$ corresponding to the singular value $\sigma_i$. 
Note that

\[ Av_i = \sigma_i u_i \quad \text{and} \quad A^* u_i = \sigma_i v_i, \quad 1 \leq i \leq p. \]

From \( AA^* = U \Sigma \Sigma^T U^* \) and \( A^* A = V \Sigma^T \Sigma V^* \) we deduce that the singular values of \( A \) are the (positive) square roots of the eigenvalues of the matrices \( AA^* \) and \( A^* A \); the left singular vectors of \( A \) are eigenvectors of \( AA^* \), and the right ones are eigenvectors of \( A^* A \).

Moreover,

\[ A = \sum_{i=1}^{r} \sigma_i u_i v_i^*, \]

showing that any matrix \( A \) of rank \( r \) is the sum of exactly \( r \) rank-one matrices.
It is also important to recall that if we form the Hermitian matrix

\[
A = \begin{bmatrix}
  0 & A \\
  A^* & 0
\end{bmatrix} \in \mathbb{C}^{(n+m) \times (n+m)},
\]

then the eigenvalues of \( A \) are of the form \( \lambda_i = \pm \sigma_i \), and the corresponding eigenvectors are given by

\[
x_i = \begin{bmatrix}
  u_i \\
  \pm v_i
\end{bmatrix}.
\]

The matrix \( A \) plays an important role in computations and in the analysis of bipartite and directed networks.
For any matrix \( A = U \Sigma V^* \in \mathbb{C}^{m \times n} \), the Moore–Penrose generalized inverse (or pseudo-inverse) of \( A \) is defined as

\[
A^\dagger := V \Sigma^\dagger U^* \in \mathbb{C}^{n \times m},
\]

where

\[
\Sigma^\dagger := \text{diag}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}, 0, \ldots, 0) \in \mathbb{R}^{n \times m}.
\]

The pseudo-inverse is the unique matrix satisfying the following four conditions:

\[
AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (AA^\dagger)^* = AA^\dagger, \quad (A^\dagger A)^* = A^\dagger A.
\]

When \( m = n \) and \( A \) is nonsingular, \( A^\dagger = A^{-1} \).
Schur Normal Form

Any matrix $A \in \mathbb{C}^{n \times n}$ is unitarily similar to an upper triangular matrix. That is, there exist $U \in \mathbb{C}^{n \times n}$ unitary and $T \in \mathbb{C}^{n \times n}$ upper triangular such that

$$U^* AU = T.$$ 

Neither $U$ nor $T$ are unique: only the diagonal elements of $T$ are, and they are the eigenvalues of $A$.

The matrix $A$ is normal if, and only if, $T$ is diagonal.

If $T$ is split as

$$T = D + N$$

with $D$ diagonal and $N$ strictly upper triangular (nilpotent), then the “size” of $N$ is a measure of how far $A$ is from normal.
The notion of a norm on a vector space (over $\mathbb{R}$ or $\mathbb{C}$) is well-known. A matrix norm on the matrix spaces $\mathbb{R}^{n \times n}$ or $\mathbb{C}^{n \times n}$ is just a vector norm $\| \cdot \|$ which satisfies the additional requirement of being *submultiplicative*:

$$\| AB \| \leq \| A \| \| B \| , \quad \forall A, B.$$  

Important examples of matrix norms include the *induced norms* (especially for $p = 1, 2, \infty$) and the *Frobenius norm*

$$\| A \|_F := \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2}.$$  

It is easy to show that

$$\| A \|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|, \quad \| A \|_\infty = \| A^* \|_1 = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$
Furthermore, denoting by $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ the singular values of $A$, it holds

$$
\|A\|_2 = \sigma_1, \quad \|A\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i^2}
$$

and therefore $\|A\|_2 \leq \|A\|_F$ for all $A$. These facts hold for rectangular matrices as well.

Also, the *spectral radius* $\rho(A) := \max\{|\lambda| : \lambda \in \sigma(A)\}$ satisfies $\rho(A) \leq \|A\|$ for all $A$ and all matrix norms.

For a normal matrix, $\rho(A) = \|A\|_2$. But if $A$ is nonnormal, $\|A\|_2 - \rho(A)$ can be arbitrarily large.

Also note that if $A$ is diagonalizable with $A = XDX^{-1}$, then

$$
\|A\|_2 = \|XDX^{-1}\|_2 \leq \|X\|_2\|X^{-1}\|_2\|D\|_2 = \kappa_2(X)\rho(A),
$$

where $\kappa_2(X) = \|X\|_2\|X^{-1}\|_2$ is the *spectral condition number* of $X$. 
Matrix powers and matrix polynomials

For a matrix $A \in \mathbb{C}^{n \times n}$, the (asymptotic) behavior of the powers $A^k$ for $k \to \infty$ is of fundamental importance.

More generally, we will be interested in studying matrix polynomials

$$p(A) = c_0I + c_1A + c_2A^2 + \cdots + c_kA^k.$$ 

Let $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, and let $A = ZJZ^{-1}$ where $J$ is the Jordan form of $A$. Then $p(A) = Zp(J)Z^{-1}$. Hence, the eigenvalues of $p(A)$ are given by $p(\lambda_i)$, for $i = 1, \ldots, n$. Moreover, $A$ and $p(A)$ have the same eigenvectors. This applies, in particular, to $p(A) = A^k$.

Thus, if $A$ is diagonalizable with $A = XDX^{-1}$ then $p(A) = Xp(D)X^{-1}$. 

We recall the following basic result concerning matrix powers:

**Theorem:** Let $A \in \mathbb{C}^{n \times n}$, then

$$\lim_{k \to \infty} A^k = 0 \iff \rho(A) < 1.$$  

**Important:** It is sometimes stated that if $\rho(A) \ll 1$, then the convergence of $A^k$ to 0 will be *very fast*. While this is true if $A$ is normal, or “close” to normal, it is *not true* in general. For highly nonnormal matrices, the convergence of $A^k$ to 0 can be *extremely slow*, even for $\rho(A) \approx 0$ (and in fact, even for $\rho(A) = 0$!).

The point is that $\rho(A)$ only governs the *asymptotic* convergence behavior of the powers of $A$, not the transient behavior; in the short term, $A^k$ can have arbitrarily large entries even if $\rho(A) \approx 0$. Moreover, the transient phase can last a long time—i.e., it may take a huge value of $k$ before the asymptotic regime kicks in.
Theorem (Cayley–Hamilton): For any matrix $A \in \mathbb{C}^{n \times n}$ it holds

$$p_A(A) = 0$$

where $p_A(\lambda) := \det(A - \lambda I)$ is the characteristic polynomial of $A$.

An even more important polynomial is the minimum polynomial of $A$, which is defined as the monic polynomial $q_A(\lambda)$ of least degree such that $q_A(A) = 0$. Note that $q_A|p_A$, hence $\deg(q_A) \leq \deg(p_A) = n$.

It easily follows from this that for any nonsingular $A \in \mathbb{C}^{n \times n}$, the inverse $A^{-1}$ can be expressed as a polynomial in $A$ of degree at most $n - 1$:

$$A^{-1} = c_0 I + c_1 A + c_2 A^2 + \cdots + c_k A^k, \quad k \leq n - 1.$$ 

Of course, the coefficients $c_i$ depend on $A$. The same result holds for powers $A^p$ with $p \geq n$, and more generally for matrix functions $f(A)$, as we will see.
Matrices and graphs

To any matrix $A \in \mathbb{C}^{n \times n}$ we can associate a directed graph, or digraph, $G(A) = (V, E)$ where $V = \{1, 2, \ldots, n\}$ and $E \subseteq V \times V$, where $(i, j) \in E$ if and only if $a_{ij} \neq 0$.

Diagonal entries are usually ignored ($\Rightarrow$ no loops in $G(A)$).

Let $|A| := (|a_{ij}|)$, then the digraph $G(|A|^2)$ is given by $(V, \hat{E})$ where $\hat{E}$ is obtained by including all directed edges $(i, k)$ such that there exists $j \in V$ with $(i, j) \in E$ and $(j, k) \in E$.

For higher powers $p$, the digraph $G(|A|^p)$ is defined similarly: its edge set consists of all pairs $(i, k)$ such that there is a directed path of length at most $p$ joining node $i$ with node $k$ in $G(A)$.

Note: the reason for the absolute value is to disregard the effect of possible cancellations in $A^p$. 


**Definition:** Let $G = (V, E)$ be a directed graph. The *transitive closure* of $G$ is the graph $\bar{G} = (V, \bar{E})$ where

$$(i, j) \in \bar{E} \iff \text{there is a directed path joining } i \text{ and } j \text{ in } G(A).$$

For Hermitian or symmetric matrices, simple (undirected) graphs can be used instead of directed graphs.

The same is true for *structurally symmetric* matrices, i.e., matrices such that $a_{ij} \neq 0 \iff a_{ji} \neq 0$.

Since most matrices arising from the discretization of PDEs are structurally symmetric, undirected graphs are most often used in this area. Also note that if $A$ is “not too far from being structurally symmetric”, then the undirected graph $G(A + A^T)$ is often used in practice.
Irreducibility

**Definition:** A matrix $A \in \mathbb{C}^{n \times n}$ is *reducible* if there exists a permutation matrix $P$ such that

$$P^T AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

with $A_{11}$ and $A_{22}$ square submatrices. If no such $P$ exists, $A$ is said to be *irreducible*.

**Theorem.** The following statements are equivalent:

(i) the matrix $A$ is irreducible
(ii) the digraph $G(A)$ is strongly connected, i.e., for every pair of nodes $i$ and $j$ in $V$ there is a directed path in $G(A)$ that starts at node $i$ and ends at node $j$
(iii) the transitive closure $\bar{G}(A)$ of $G(A)$ is the *complete graph* on $V$, i.e., the graph with edge set $E = V \times V$.

Note that (iii) and the Cayley–Hamilton Theorem imply that the powers $(I + A)^p$ are completely full for $p \geq n - 1$ (barring cancellation).
**Definition:** A matrix $A \in \mathbb{R}^{n \times n}$ is *nonnegative* if all its entries are nonnegative: $a_{ij} \geq 0$ for all $i, j = 1, \ldots, n$.

Also, $A$ is *positive* if $a_{ij} > 0$ for all $i, j = 1, \ldots, n$.

We write $A \geq 0$ and $A > 0$, respectively.

**Note:** $A \geq 0$ if and only if $A(\mathbb{R}_+) \subseteq \mathbb{R}_+$, where

\[
\mathbb{R}_+^n := \{x = (x)_i \in \mathbb{R}^n \mid x_i \geq 0, \forall i = 1, \ldots, n\}.
\]

**Theorem:** An $n \times n$ nonnegative matrix $A$ is irreducible if and only if $(I + A)^k > 0$ for some positive integer $k \leq n - 1$. 
**Theorem (Perron–Frobenius):** Let $A \geq 0$ be irreducible, then

(i) $\rho(A)$ is a simple eigenvalue of $A$;

(ii) There exists a positive eigenvector $x = (x_i)$ associated with $\rho(A)$:

$$Ax = \rho(A)x, \quad x_i > 0 \quad \forall i = 1, \ldots, n;$$

(iii) $\rho(A)$ increases when any entry of $A$ increases; that is,

$$\tilde{A} \geq A \quad \text{and} \quad \tilde{A} \neq A \quad \Rightarrow \quad \rho(\tilde{A}) > \rho(A)$$

Here $\tilde{A} \geq A$ means that $\tilde{A} - A$ is nonnegative.

The Perron–Frobenius Theorem holds in particular for positive matrices.
Oskar Perron (1880-1975) and Georg Ferdinand Frobenius (1849-1917)
**Definition:** A matrix $A \in \mathbb{R}^{n \times n}$ is an *M-matrix* if it can be written as

$$A = rI - B,$$

where $B \geq 0$ and $r \geq \rho(B)$.

If $r > \rho(B)$ then $A$ is a *nonsingular M-matrix*.

**Theorem:** A matrix $A \in \mathbb{R}^{n \times n}$ is a nonsingular $M$-matrix if and only if $a_{ij} \leq 0$ for all $i \neq j$ and $A$ is invertible with $A^{-1} \geq 0$.

**Note:** If $A$ is an *irreducible* nonsingular $M$-matrix, then $A^{-1} > 0$.

**Definition:** An invertible matrix $A \in \mathbb{R}^{n \times n}$ is *monotone* if $A^{-1} \geq 0$.

- Matrix monotonicity corresponds to a *discrete maximum principle*.
- All nonsingular $M$-matrices are monotone, but the converse is not true.
**Theorem:** Let $A \in \mathbb{R}^{n \times n}$ be a symmetric nonsingular $M$-matrix. Then $A$ is positive definite.

Symmetric nonsingular $M$-matrices are sometimes called *Stieltjes matrices*.

A nonsymmetric nonsingular $M$-matrix is not necessarily positive definite; that is, $A + A^T$ may be indefinite. A simple example is given by

$$A = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$$

However, all nonsingular $M$-matrices are *positive stable*: $Re(\lambda) > 0$, for all $\lambda \in \sigma(A)$.

$M$-matrices are extremely well-behaved from the numerical point of view.
Geršgorin’s Theorem

Let $A \in \mathbb{C}^{n \times n}$. For all $i = 1 : n$, let

$$r_i := \sum_{j \neq i} |a_{ij}|, \quad D_i = D_i(a_{ii}, r_i) := \{ z \in \mathbb{C} : |z - a_{ii}| \leq r_i \}.$$ 

The set $D_i$ is called the $i$th Geršgorin disk of $A$.

Geršgorin’s Theorem (1931) states that $\sigma(A) \subset \bigcup_{i=1}^{n} D_i$. Moreover, each connected component of $\bigcup_{i=1}^{n} D_i$ consisting of $p$ Geršgorin disks of $A$ contains exactly $p$ eigenvalues of $A$, counted with their multiplicities.

Of course, the same result holds replacing the off-diagonal row-sums with off-diagonal column-sums. The spectrum is then contained in the intersection of the two resulting regions.
The field of values (or numerical range) of $A \in \mathbb{C}^{n \times n}$ is the set

$$\mathcal{W}(A) := \{ z = \langle Ax, x \rangle \mid x^* x = 1 \}.$$ 

This set is a compact subset of $\mathbb{C}$ containing the eigenvalues of $A$; it is also convex. This last statement is known as the Hausdorff–Toeplitz Theorem, and is highly nontrivial.

The definition of numerical range also applies to bounded linear operators on a Hilbert space $\mathcal{H}$; however, $\mathcal{W}(A)$ may not be closed if $\dim(\mathcal{H}) = \infty$.

Field of Values of a random $10 \times 10$ matrix.
Outline

1 Plan of these lectures
2 Review of matrix theory
3 Functions of matrices
4 Bibliography
Motivation

Both classical and new applications have resulted in increased interest in the theory and computation of matrix functions over the past few years:

- Solution of time-dependent ODEs/PDEs
- Quantum chemistry (electronic structure theory)
- Applications to Network Science
- Theoretical particle physics (QCD)
- Markov models in finance
- Data mining
- Control theory
- etc.

Currently a hot topic in scientific computing!
A bit of history

Simple matrix functions appear already in Cayley’s *A memoir on the theory of matrices* (1858). This is considered to be “the first paper to investigate the algebraic properties of matrices regarded as objects of study in their own right” (N. J. Higham).

The term *matrix* itself had been introduced by Sylvester already in 1850.

In his paper Cayley considered square roots of matrices, as well as polynomial and rational functions of a matrix (the simplest of which is, of course, $A^{-1}$). The paper also contains a statement of the Cayley–Hamilton Theorem.

Founding fathers

Arthur Cayley (1821-1895) and James Joseph Sylvester (1814-1897)
The first general definitions of matrix function begin to appear after 1880. A completely satisfactory definition, however, will have to wait until 1932.

There have been proposed in the literature since 1880 eight distinct definitions of a matric function, by Weyr, Sylvester and Buchheim, Giorgi, Cartan, Fantappié, Cipolla, Schwerdtfeger and Richter [...] All of the definitions except those of Weyr and Cipolla are essentially equivalent.

Matrix function as defined by Sylvester (1883) and Buchheim (1886)

Polynomial interpolation

Let $\lambda_1, \ldots, \lambda_s$ be the distinct eigenvalues of $A \in \mathbb{C}^{n \times n}$ and let $n_i$ be the index of $\lambda_i$. Then $f(A) := r(A)$, where $r$ is the unique Lagrange–Hermite interpolating polynomial of degree $< \sum_{i=1}^{s} n_i$ satisfying

$$r^{(j)}(\lambda_i) = f^{(j)}(\lambda_i) \quad j = 0, \ldots, n_i - 1, \quad i = 1, \ldots, s.$$ 

Of course, this implies that the values $f^{(j)}(\lambda_i)$ with $0 \leq j \leq n_i - 1$ and $1 \leq i \leq s$ exist. We say that $f$ is defined on the spectrum of $A$. When all the eigenvalues are distinct, the interpolation polynomial has degree $n - 1$.

Remark:
Every matrix function is a polynomial in $A$!
Matrix function as defined by Weyr (1887)

Taylor series

Suppose $f$ has a Taylor series expansion

$$f(z) = \sum_{k=0}^{\infty} a_k (z - \alpha)^k \quad \left( a_k = \frac{f^{(k)}(\alpha)}{k!} \right)$$

with radius of convergence $r$. If $A \in \mathbb{C}^{n \times n}$ and each of the distinct eigenvalues $\lambda_1, \ldots, \lambda_s$ of $A$ satisfies

$$|\lambda_i - \alpha| < r,$$

then

$$f(A) := \sum_{k=0}^{\infty} a_k (A - \alpha I)^k.$$
Matrix function as defined by Giorgi (1928)

**Jordan canonical form**

Let $A \in \mathbb{C}^{n \times n}$ have Jordan canonical form $Z^{-1}AZ = J$ with $J = \text{diag}(J_1, \ldots, J_p)$. We define

$$f(A) := Z f(J) Z^{-1} = Z \text{diag}(f(J_k(\lambda_k))) Z^{-1},$$

where

$$f(J_k(\lambda_k)) = \begin{pmatrix} f(\lambda_k) & f'(\lambda_k) & \cdots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ f(\lambda_k) & \ddots & \ddots & \vdots \\ \vdots & \ddots & f'(\lambda_k) & f(\lambda_k) \end{pmatrix}.$$

**Remark:** If $A = XD X^{-1}$ with $D$ diagonal, then

$$f(A) := X f(D) X^{-1} = X \text{diag}(f(\lambda_i)) X^{-1}.$$
Matrix function as defined by E. Cartan (1928)

In a letter to Giovanni Giorgi, Cartan proposed the following definition:

**Contour integral**

Let $f$ be analytic inside a closed simple contour $\Gamma$ enclosing $\sigma(A)$, the spectrum of $A$. Then

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

where the integral is taken entry-wise.

**Remarks:** The contour integral approach to $f(A)$ had already been used in special cases by Frobenius (1896) and by Poincaré (1899).

This definition can also be used to define analytic functions of operators, and more generally analytic functions over Banach algebras ("holomorphic functional calculus"). We will return to this topic in Lecture IV.
Two pioneers of matrix functions

Eduard Weyr (1852-1903) and Élie Cartan (1869-1951)
Giorgi’s definition assumes that whenever \( f \) is a multi-valued function (e.g., the square root or the logarithm), the same branch of \( f \) is used for every Jordan block of \( A \). Such matrix functions are called *primary matrix functions*.

Functions that allow using different branches of \( f \) for different Jordan blocks are called *non-primary*. The most general definition of a matrix function, due to Cipolla (1932), includes non-primary functions.

For example,

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
-1 & 0 \\
0 & -1
\end{bmatrix}
\]

are both primary square roots of \( I_2 \), but

\[
\begin{bmatrix}
1 & 0 \\
0 & -1
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
-1 & 0 \\
0 & 1
\end{bmatrix}
\]

are non-primary. Here we only consider primary matrix functions.
Early Italian contributors to matrix functions

Giovanni Giorgi (1871-1950) and Michele Cipolla (1880-1947)
Giorgi introduced a precursor of the metric system based on four units.
And if you are really interested...


Some basic facts about matrix functions

Let $A \in \mathbb{C}^{n \times n}$ and let $f$ be defined on $\sigma(A)$, then

- $f(A)A = Af(A)$;
- $f(A^T) = f(A)^T$;
- $f(XAX^{-1}) = Xf(A)X^{-1}$;
- $\sigma(f(A)) = f(\sigma(A))$;
- $(\lambda, x)$ eigenpair of $A \Rightarrow (f(\lambda), x)$ eigenpair of $f(A)$;
- If $A = (A_{ij})$ is block triangular then $F = f(A)$ is block triangular with the same block structure as $A$, and $F_{ii} = f(A_{ii})$;
- $f(\text{diag}(A_{11}, \ldots, A_{pp})) = \text{diag}(f(A_{11}), \ldots, f(A_{pp}))$;
- $f(I_m \otimes A) = I_m \otimes f(A)$, where $\otimes$ is the Kronecker product;
- $f(A \otimes I_m) = f(A) \otimes I_m$.

For proofs, see Higham (2008).
Some basic facts about matrix functions (cont.)

**Theorem** (Higham, Mackey, Mackey, and Tisseur): Let \( f \) be analytic on an open set \( \Omega \subseteq \mathbb{C} \) such that each connected component of \( \Omega \) is closed under conjugation. Consider the corresponding matrix function \( f \) on the set \( \mathcal{D} = \{ A \in \mathbb{C}^{n \times n} : \sigma(A) \subseteq \Omega \} \). Then the following are equivalent:

(a) \( f(A^*) = f(A)^* \) for all \( A \in \mathcal{D} \).
(b) \( f(A) = f(\overline{A}) \) for all \( A \in \mathcal{D} \).
(c) \( f(\mathbb{R}^{n \times n} \cap \mathcal{D}) \subseteq \mathbb{R}^{n \times n} \).
(d) \( f(\mathbb{R} \cap \Omega) \subseteq \mathbb{R} \).

In particular, if \( f(x) \in \mathbb{R} \) for \( x \in \mathbb{R} \) and \( A \) is Hermitian, so is \( f(A) \).
Let $A \in \mathbb{C}^{n \times n}$ and let $f$ be defined on $\sigma(A)$. The following expressions (in increasing order of generality) are often useful:

- If $A$ is normal (in particular, Hermitian) then
  \[ f(A) = \sum_{i=1}^{n} f(\lambda_i) u_i u_i^* \]

- If $A \in \mathbb{C}^{n \times n}$ is diagonalizable then
  \[ f(A) = \sum_{i=1}^{n} f(\lambda_i) x_i y_i^* \]

- If $A$ is arbitrary then
  \[ f(A) = \sum_{i=1}^{s} \sum_{j=0}^{n_i-1} \frac{f(j)(\lambda_i)}{j!} (A - \lambda_i I)^j G_i \]
An expression for $f(A)$ can also be obtained from the Schur form of $A$, $A = U T U^*$ with $T = (t_{ij})$ upper triangular:

$$f(A) = U f(T) U^*, \quad f(T) = (f_{ij})$$

where $f_{ij} = 0$ for $i > j$, $f_{ij} = f(\lambda_i)$ for $i = j$, and

$$f_{ij} = \sum_{(s_0, \ldots, s_k) \in S_{ij}} t_{s_0, s_1} t_{s_1, s_2} \cdots t_{s_{k-1}, s_k} f[\lambda_{s_0}, \ldots, \lambda_{s_k}] \quad \text{for} \quad i < j.$$  

Here $S_{ij}$ is the set of all strictly increasing sequences of integers starting at $i$ and ending at $j$, and $f[\lambda_{s_0}, \ldots, \lambda_{s_k}]$ is the $k$th order divided difference of $f$ at $\{\lambda_{s_0}, \ldots, \lambda_{s_k}\}$.
Important examples

Let $A \in \mathbb{C}^{n \times n}$, and let $z \notin \sigma(A)$. The resolvent of $A$ at $z$ is defined as

$$R(A; z) = (zI - A)^{-1}.$$

The resolvent is central to the definition of matrix functions via the contour integral approach. As a special case, the resolvent can be used to define the spectral projector onto the eigenspace of a matrix or operator corresponding to an isolated eigenvalue $\lambda_0 \in \sigma(A)$:

$$P_{\lambda_0} := \frac{1}{2\pi i} \int_{|z-\lambda_0|=\varepsilon} (zI - A)^{-1}dz$$

where $\varepsilon > 0$ is small enough so that no other eigenvalue of $A$ falls within $\varepsilon$ of $\lambda_0$.

Remarks: More generally, one can define the spectral projector onto the invariant subspace of $A$ corresponding to a set of selected eigenvalues by integrating $R(A; z)$ along a contour surrounding those eigenvalues and excluding the others. The spectral projector is orthogonal if and only if $A$ is normal.
Also extremely important is the matrix exponential

\[ e^A = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} A^k \]

which is defined for arbitrary \( A \in \mathbb{C}^{n \times n} \).

Just as the resolvent is central to spectral theory, the matrix exponential is fundamental to the solution of differential equations. For example, the solution to the inhomogeneous system

\[ \frac{dy}{dt} = Ay + f(t, y), \quad y(0) = y_0, \quad y \in \mathbb{C}^n, \quad A \in \mathbb{C}^{n \times n} \]

is given (implicitly!) by

\[ y(t) = e^{tA} y_0 + \int_0^t e^{A(t-s)} f(s, y(s)) ds. \]

In particular, \( y(t) = e^{tA} y_0 \) when \( f = 0 \).

It is important to recall that \( \lim_{t \to \infty} e^{tA} = 0 \) if and only if \( A \) is a stable matrix: \( \Re(\lambda) < 0 \), for all \( \lambda \in \sigma(A) \).
Important examples (cont.)

When \( f(t, y) = b \in \mathbb{C}^n \) (=const.), the solution can also be expressed as

\[
y(t) = t \psi_1(tA)(b + Ay_0) + y_0
\]

where

\[
\psi_1(z) = \frac{e^z - 1}{z} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots
\]

Trigonometric functions and square roots of matrices are also important in applications. For example, the solution to the second-order system

\[
\frac{d^2y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y'_0
\]

(where \( A \) is SPD) can be expressed as

\[
y(t) = \cos(\sqrt{A}t) y_0 + (\sqrt{A})^{-1} \sin(\sqrt{A}t) y'_0.
\]
Important examples (cont.)

If $A, B \in \mathbb{C}^{n \times n}$ commute ($AB = BA$), then the identity

$$e^{A+B} = e^A e^B$$

holds true, but in general $e^{A+B} \neq e^A e^B$.

A useful identity that is always true is

$$e^A = \cosh(A) + \sinh(A),$$

where

$$\cosh(A) = \frac{e^A + e^{-A}}{2} = \sum_{k=0}^{\infty} \frac{1}{(2k)!} A^{2k}$$

and

$$\sinh(A) = \frac{e^A - e^{-A}}{2} = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} A^{2k+1}.$$
Apart from the contour integration formula, the matrix exponential and the resolvent are also related through the *Laplace transform*: there exists an $\omega \in \mathbb{R}$ such that $z \notin \sigma(A)$ for $\text{Re}(z) > \omega$ and

$$(z I - A)^{-1} = \int_0^\infty e^{-zt} e^{tA} \, dt = \int_0^\infty e^{-t(zI - A)} \, dt.$$ 

Recall also that if $|z| > \rho(A)$, the following *Neumann series expansion* of the resolvent is valid:

$$(z I - A)^{-1} = z^{-1}(I + z^{-1}A + z^{-2}A^2 + \cdots) = z^{-1} \sum_{k=0}^{\infty} z^{-k} A^k.$$ 

A useful variant of this expression is

$$(I - \alpha A)^{-1} = I + \alpha A + \alpha^2 A^2 + \cdots = \sum_{k=0}^{\infty} \alpha^k A^k,$$

where $0 < \alpha < \frac{1}{\rho(A)}$. 

Important examples (cont.)
Outline

1. Plan of these lectures
2. Review of matrix theory
3. Functions of matrices
4. Bibliography


