Exploiting Localization in Matrix Computations
I. Motivation and Background

Michele Benzi

Department of Mathematics and Computer Science

Emory University

Atlanta, Georgia, USA

CIME-EMS Summer School in Applied Mathematics

Exploiting Hidden Structure in Matrix Computations:
Theory and Applications

Cetraro, 22-26 June, 2015

Outline

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

Outline

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

Plan and acknowledgements

- Lecture I: Motivation. Linear algebra background.
- Lecture II: Functions of matrices.
- Lecture III: Localization in matrix functions.
- Lecture IV: Locality in electronic structure computations.

Many thanks to the CIME Foundation, to GNCS, to the EMS and to my friends and colleagues Paola Boito, Nader Razouk, and Valeria Simoncini. 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)

Outline

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

Locality in physics and in mathematics

In physics, the term localization is often used to describe two types of situations:

1. A function decays rapidly to zero outside of a small region: the function is localized in space.
2. The interaction strength between the different parts of a system extended in space decreases rapidly with the distance: correlations are short-ranged.

The opposite of localization is delocalization: a function is delocalized if it's non-negligible on an extended region.

Similarly, if non-local (long-range) interactions are important, a system is delocalized.

Locality in physics and in mathematics (cont.)

In quantum mechanics, the stationary states of a system (e.g., a particle) are described by wave functions, $\Psi_n(\mathbf{r})$, $n = 0, 1, \ldots$ The probability that a particle with wave function Ψ_n is found in a given region $\Omega \subseteq \mathbf{R}^3$ is given by

$$\mathbf{Pr}\,(\text{particle in}\,\Omega) = \int_\Omega |\Psi_n(\mathbf{r})|^2\,\mathrm{d}\mathbf{r}.$$

As an example, consider the electron in a hydrogen atom. The radial part of the first atomic orbital, the wave function corresponding to the lowest energy (ground state), is a decaying exponential:

$$\Psi_0(r) = \frac{1}{\sqrt{\pi}\,a_0^{3/2}}\,\mathrm{e}^{-r/a_0}, \quad r \geq a_0,$$

where $a_0 = \frac{\hbar^2}{me^2} = 0.0529$ nm is the Bohr radius. Thus, the wave function is strongly localized in space.

Locality in physics and in mathematics (cont.)

Localization of the wave function Ψ_0 expresses the fact that in the hydrogen atom at ground state, the electron is bound to a small region around the nucleus, and the probability of finding the electron at a distance r decreases rapidly as r increases.

The wave function Ψ_0 satisfies the (stationary) Schrödinger equation:

$$H\,\Psi_0 = E_0\,\Psi_0$$

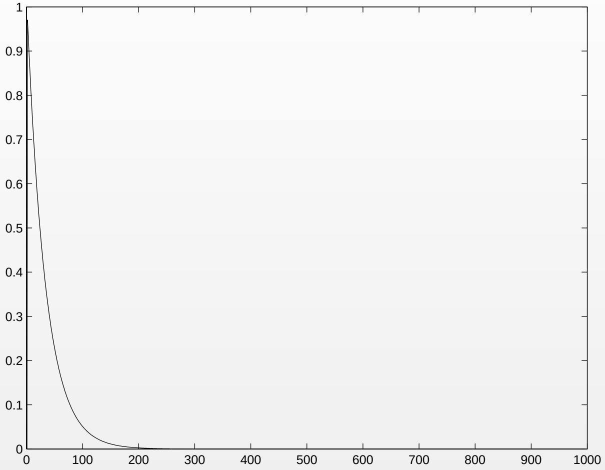
where (using atomic units)

$$H = -\frac{1}{2}\Delta - \frac{1}{r} \qquad (r = \sqrt{x^2 + y^2 + z^2}\,)$$

is the Hamiltonian, or energy, operator, and E_0 is the ground state energy.

That is, the ground state Ψ_0 is the eigenfunction of the Hamiltonian corresponding to the lowest eigenvalue E_0.

Locality in physics and in mathematics (cont.)



Electron in a Coulomb field: the wave function is localized in space.

Locality in physics and in mathematics (cont.)

Note that the Hamiltonian is of the form $H = T + V$ where

$$T = -\frac{1}{2}\Delta = \text{kinetic energy}$$

and

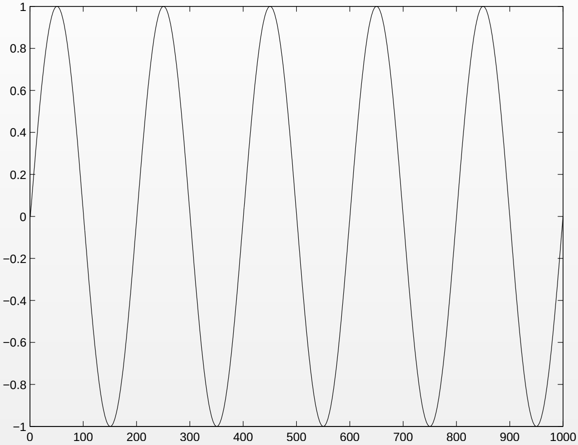
$$V = -\frac{1}{r} = \text{(Coulomb) potential.}$$

What happens if the Coulomb potential is absent? In this case there is no force binding the electron to the nucleus: the electron is “free".

This implies the delocalization of the corresponding wave function.

Consider for semplicity the 1D case: if we confine the system to the interval $[0, L]$ (with zero Dirichlet boundary conditions), then the eigenfunction corresponding to the smallest eigenvalue of the Hamiltonian $H = -\frac{d^2}{dx^2}$ is $\Psi_0(x) = \sin\left(\frac{2\pi}{L}x\right)$, which is delocalized. Similarly in 2D and 3D.

Locality in physics and in mathematics (cont.)



Electron in a box: the wave function is delocalized.

Locality in physics and in mathematics (cont.)

Consider now an extended system consisting of a large number of atoms, assumed to be in the ground state.

Suppose the system is perturbed at one point space, for example by hitting it with a small amount of radiation.

If the system is an insulator, then the effect of the perturbation will only be felt locally: it will not be felt outside of a small region. This “absence of diffusion" is also known as localization.

W. Kohn called this the “nearsightedness" of electronic matter. In insulators, and also in semi-conductors under suitable conditions, the electrons tend to stay put.

Locality in physics and in mathematics (cont.)

In a metallic system, in contrast, local perturbations can have long-range effects. The electrons are free to move around, and the system's electron density is delocalized.

Localization is a phenomenon of major importance in quantum chemistry and in solid state physics. We will return on this in our last lecture, on the electronic structure problem.

Locality (or lack thereof) is also of central importance in quantum information theory and quantum computing, in connection with the notion of entanglement of states.

J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, Rev. Modern Phys., 82 (2010), pp. 277–306.

Locality in physics and in mathematics (cont.)

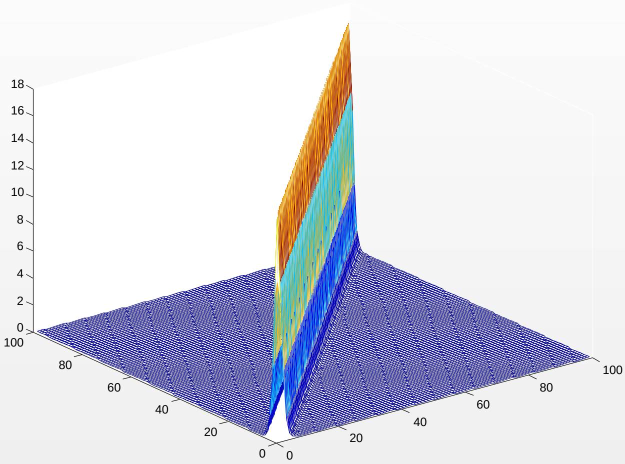
The study of localization in numerical mathematics is a more recent phenomenon.

It emerged in the 1980s as a results of various trends in numerical analysis, in particular in the study of wavelets and in problems of approximation theory and in numerical linear algebra. Its importance has been rapidly increasing in recent years.

Locality in numerical linear algebra is related to, but should not be confused with, sparsity.

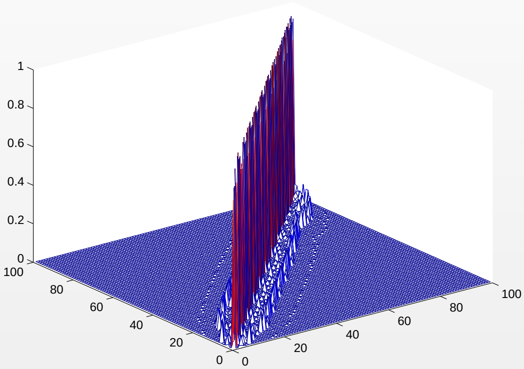
For example: a matrix can be localized even if it is not sparse, although it will be close to a sparse matrix (in a suitable norm).

Locality in physics and in mathematics (cont.)



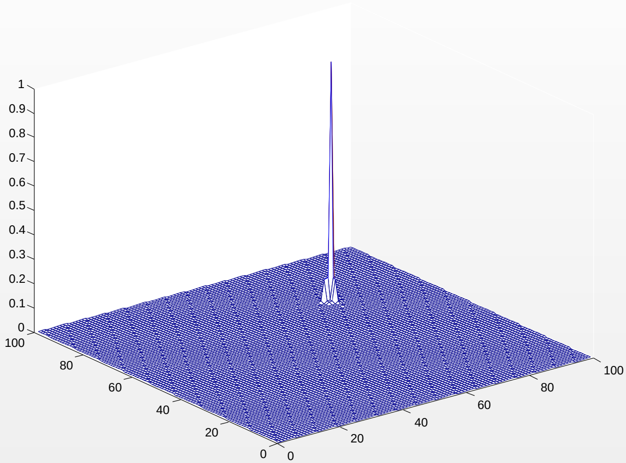
The exponential of a tridiagonal matrix (discrete 1D Laplacian).

Locality in physics and in mathematics (cont.)



The square root of a sparse matrix (`nos4` from Matrix Market).

Locality in physics and in mathematics (cont.)



The spectral projector onto the invariant subspace corresponding to an isolated eigenvalue of a banded matrix.

Locality in physics and in mathematics (cont.)

Perhaps less obviously, a (discrete) system could well be described by a highly sparse matrix but be strongly delocalized.

Think of a small-world network, like Facebook, or an expander graph.

Even if, on average, each component of such a system is directly connected to only a few other components, the system is strongly delocalized, since every node is only a few steps away from every other node.

Hence, a “disturbance" at one node propagates quickly to the entire system. Every short range interaction is also long-range: localization is impossible in such systems.

This can be a very valuable property in certain cases!

Locality in physics and in mathematics (cont.)

Exploiting locality can lead to important speed-ups and even make apparently intractable problems tractable.

This is especially true for problems involving matrix functions, but localization in eigenvectors and in the solution of sparse linear systems is also very useful, when present.

The theory extends to matrices whose elements are not scalars but may be functions, or other matrices, or even operators on an infinite-dimensional Hilbert space. This has applications in various parts of physics, not just in numerical analysis.

As an illustration of the theory, in the last lecture I will go back to quantum mechanics and discuss the importance of locality in the computation of electronic structures.

Locality in physics and in mathematics (cont.)

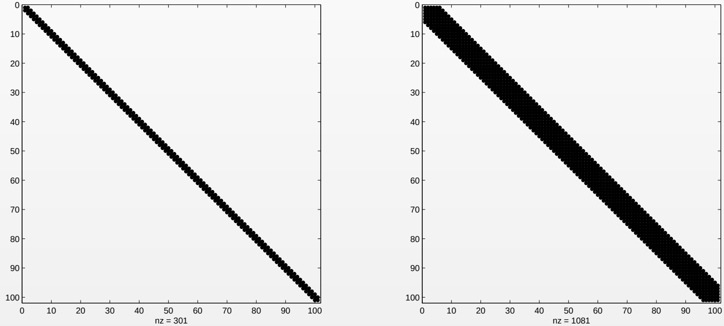
Interestingly, delocalization can also be computationally advantageous.

Consider for example the iterative solution of linear systems, or eigenvalue problems. The fact that “information" spreads very quickly across the computational “domain" typically means fast convergence of even simple iterative methods, independent of system size.

In contrast, in a highly localized system information tends do diffuse slowly, and many iterations are required for convergence.

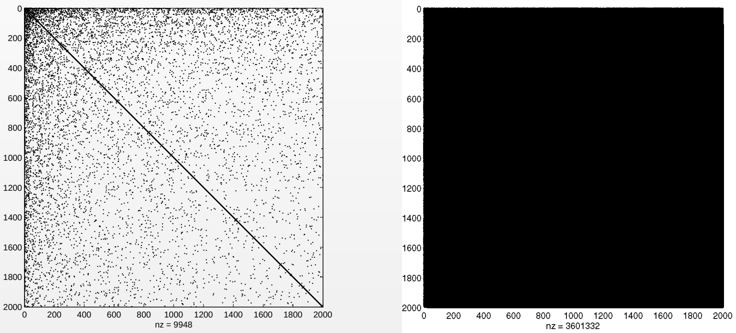
On the other hand, delocalized systems are much more tightly coupled than localized one, which makes parallelization a very hard task.

Locality in physics and in mathematics (cont.)



A localized system: 1D graph Laplacian L (left) and the matrix L^5 (right).

Locality in physics and in mathematics (cont.)



Small-world network: graph Laplacian L (left) and the matrix L^5 (right).

Locality in physics and in mathematics (cont.)

Note: Things I will **not** cover include

- Sparse solutions to dense linear algebra problems (compressed sensing, ℓ_1 minimization, etc.)
- Data sparse approximations of non-local operators (hierarchical, semi-separable, quasi-separable matrices, and related structures)
- Applications to signal processing, computational harmonic analysis, quantum computing, etc.

Outline

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

Matrix classes

We will be dealing primarily with matrices with entries in $\mathbb{R}$ or $\mathbb{C}$.

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

$$Z^{-1}AZ = J = \mathrm{diag}\,(J_1, J_2, \ldots, J_p),$$

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{n_k\times n_k},$$

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 λ_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 λ_i appears is called the *index* of λ_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 1×1.

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:

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

where $\lambda_1, \ldots, \lambda_s$ are the distinct eigenvalues of A, G_i is the projector onto the generalized eigenspace Ker$((A - \lambda_i I)^{n_i})$ along Ran$((A - \lambda_i I)^{n_i})$ with $n_i =$ index(λ_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 λ_i. Hence, A is a weighted sum of at most n rank-one matrices (oblique projectors).

Some useful expressions (cont.)

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 λ_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

$$\mathsf{Tr}(A) := \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$$

and, more generally,

$$\mathsf{Tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k\,, \quad \forall k = 1, 2, \ldots$$

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

$$U^* A V = \Sigma = \mathrm{diag}\,(\sigma_1, \ldots, \sigma_p)$$

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

The σ_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 $r = \mathrm{rank}(A)$. The matrix Σ 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 σ_i.

Singular Value Decomposition (cont.)

Note that

$$Av_i = \sigma_i u_i \quad \text{and} \quad A^* u_i = \sigma_i v_i, \quad 1 \le i \le p.$$

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.

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

$$U^* A U = T\,.$$

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.

Matrix norms

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*:

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

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

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|, \quad \|A\|_\infty = \|A^*\|_1 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|\,.$$

Matrix norms (cont.)

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 $A \in \mathbb{C}^{n \times n}$ and a scalar polynomial

$$p(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_k\lambda^k,$$

define

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots + c_k A^k \,.$$

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 $p(A) = Xp(D)X^{-1}$.

Matrix powers and matrix polynomials (cont.)

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 $\deg(q_A) \le \deg(p_A) = n$.

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 $n-1$:

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

Of course, the coefficients c_i depend on A. The same result holds for powers A^p with $p \ge n$, 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.

Matrices and graphs (cont.)

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

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

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 $a_{ij} \neq 0 \Leftrightarrow a_{ji} \neq 0$.

Since most matrices arising form 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

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

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).

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 *ith Geršgorin disk* of A.

Geršgorin's Theorem (1931) states that $\sigma(A) \subset \cup_{i=1}^{n} D_i$. Moreover, each connected component of $\cup_{i=1}^{n} D_i$ 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

The *field of values* (or *numerical range*) of $A \in \mathbb{C}^{n\times n}$ is the set

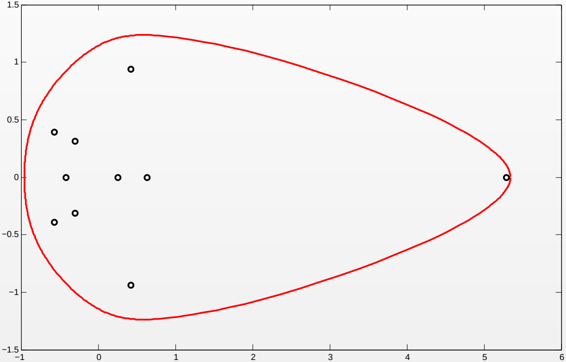
$$\mathcal{W}(A) := \{z = \langle Ax, x\rangle \,|\, x^*x = 1\}\,.$$

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 $\dim(\mathcal{H}) = \infty$.

R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1994.

The Field of Values (cont.)



Field of Values of a random 10×10 matrix.

Outline

1. Plan of these lectures
2. Motivation
3. Review of matrix theory
4. **Functions of matrices**
5. 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)
- 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

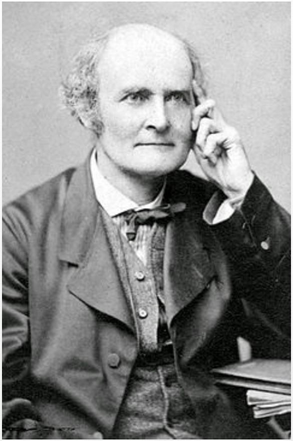
Simple matrix functions appear already in Cayley's *A memoir on the theory of matrices* (1858). This is considered to be *"the first paper to investigate the algebraic properties of matrices regarded as objects of study in their own right"* (N. J. Higham).

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.

A. Cayley, *A memoir on the theory of matrices*, Phil. Trans. Roy. Soc. London, 148:17–37, 1858.

Founding fathers



Arthur Cayley (1821-1895) and James Joseph Sylvester (1814-1897)

Definitions of Matrix Function

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.*

R. F. Rinehart, *The equivalence of definitions of a matric function*, Amer. Math. Monthly, 62:395–414, 1955.

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 λ_i. Then $f(A) := r(A)$, where r is the unique Lagrange–Hermite interpolating polynomial of degree $< \sum_{i=1}^{s} n_i$ satisfying

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

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

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

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 $J = \operatorname{diag}(J_1, \ldots, J_p)$. We define

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

where

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

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

$$f(A) := Xf(D)X^{-1} = X\operatorname{diag}(f(\lambda_i))X^{-1}.$$

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 Γ enclosing $\sigma(A)$, the spectrum of A. Then

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

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").

Two pioneers of matrix functions

Eduard Weyr (1852-1903) and Élie Cartan (1869-1951)

Primary vs. non-primary matrix functions

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)

Commemorative stamp of Giovanni Giorgi (1990)

SISTEMA GIORGI DI UNITA' DI MISURA CON 4 UNITA' DI BASE
IEC=CEI 1935-1938 GIOVANNI GIORGI 1871-1950
MKS
ITALIA 600
I.P.Z.S. - ROMA - 1990
M.M. TUCCELLI

Giorgi introduced a precursor of the metric system based on four units.

Some basic facts about matrix functions

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

- $f(A)A = Af(A)$;
- $f(A^T) = f(A)^T$;
- $f(XAX^{-1}) = Xf(A)X^{-1}$;
- $\sigma(f(A)) = f(\sigma(A))$;
- (λ, 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 $F_{ii} = f(A_{ii})$;
- $f(\mathrm{diag}\,(A_{11}, \dots, A_{pp})) = \mathrm{diag}\,(f(A_{11}), \dots, f(A_{pp}))$;
- $f(I_m \otimes A) = I_m \otimes f(A)$, 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 Ω 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) $f(\overline{A}) = \overline{f(A)}$ 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 $f(A)$.

Some basic facts about matrix functions (cont.)

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

$$f(A) = \sum_{i=1}^{s} \sum_{j=0}^{n_i-1} \frac{f^{(j)}(\lambda_i)}{j!}\,(A-\lambda_i I)^j G_i$$

Some basic facts about matrix functions (cont.)

An expression for $f(A)$ can also be obtained from the Schur form of A, $A = UTU^*$ with $T = (t_{ij})$ upper triangular:

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

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

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

Here S_{ij} is the set of all strictly increasing sequences of integers starting at i and ending at j, and $f[\lambda_{s_0},\dots,\lambda_{s_k}]$ is the kth order divided difference of f at $\{\lambda_{s_0},\dots,\lambda_{s_k}\}$. The triangular matrix function $f(T)$ can be computed by the Schur–Parlett algorithm.

G. H. Golub & C. F. Van Loan, *Matrix Computations. Fourth Edition*, Johns Hopkins University Press, 2013.

Outline

1. Plan of these lectures

2. Motivation

3. Review of matrix theory

4. Functions of matrices

5. Bibliography

Bibliography

F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 2 Volls., 1959.

G. H. Golub & C. F. Van Loan, *Matrix Computations. Fourth Edition*, Johns Hopkins University Press, 2013.

N. J. Higham, *Functions of Matrices. Theory and Computation*, SIAM, Philadelphia, 2008.

R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1991.

R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1994.

C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

L. Salce, *Lezioni sulle Matrici*, Decibel/Zanichelli, Padova/Bologna, 1993.

And if you are really interested...

J. J. Sylvester, *On the equation to the secular inequalities in the planetary theory*, Phil. Mag., 16: 267–269, 1883.

A. Buchheim, *On the theory of matrices*, Proc. London Math. Soc., 16:63–82, 1884.

—, *An extension of a theorem of Professor Sylvester's relating to matrices*, Phil. Mag., 22: 173–174, 1886. Fifth Series.

E. Weyr, *Note sur la théorie des quantités complexes formées avec n unités principales*, Bull. Sci. Math. II, 11:205–215, 1887.

G. Giorgi, *Nuove osservazioni sulle funzioni delle matrici*, Atti Accad. Lincei Rend., 6(8):3–8, 1928.

M. Cipolla, *Sulle matrici espressioni analitiche di un'altra*, Rend. Circolo Matematico Palermo, 56:144–154, 1932.