NTNU
Norwegian University of
Science and Technology

Numerical Linear Algebra

Bawfeh K. Kometa

TMA4205
August 22, 2012

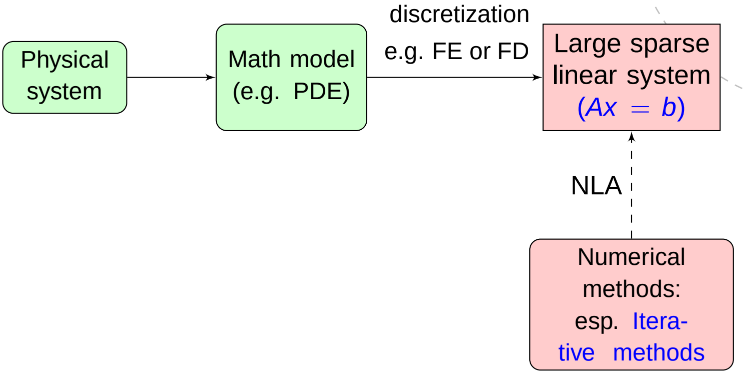
Outline

Introduction

- Motivation
- Vector norms
- Matrix norms
- Matrix-vector products
- Some important classes of matrices
- Eigenvalues and Eigenvectors
- Some matrix decompositions (or factorizations)
- Similarity transforms

Iterative methods

Motivation



Here $x \in \mathbb{R}^n$ is the unknown, while $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$ are given.

Vector norms

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$

$$\|x\|_1 = \sum_{j=1}^{n} |x_j| \qquad \text{(1-norm)}$$

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} \qquad \text{(2-norm)}$$

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} \qquad \text{(p-norm)}$$

$$\|x\|_\infty = \max_{1 \le j \le n} |x_j| \qquad \text{(max-norm)}$$

Vector norms

Inequalities

- $\|x\| \geq 0, \quad \forall x \in \mathbb{C}^n$
- $\|\alpha x\| = |\alpha| \, \|x\| \, , \quad \forall x \in \mathbb{C}^n, \, \forall \alpha \in \mathbb{C}$
- $\|x + y\| \leq \|x\| + \|y\| \, , \quad \forall x, y \in \mathbb{C}^n$ (triangle inequality)
- $|(x, y)| \leq \|x\| \, \|y\| \, ,$ (Cauchy-Schwarz)

where (x, y) denote the inner-product associated to $\|\cdot\|$

NTNU
Norwegian University of
Science and Technology

VECTOR NORMS

Euclidean inner-product

Let $x, y \in \mathbb{C}^n$.

Write $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

$$(x, y) = x^* y = \begin{bmatrix} \bar{x}_1, \ldots, \bar{x}_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{j=1}^{n} \bar{x}_j y_j$$

where x^* denotes the *adjoint*. In the real case (i.e. when $x \in \mathbb{R}^n$) we write x^* as x^T (the *transpose*).

[1]Observe that $\|x\|_2 = \sqrt{x^* x}$

---

[1] $\|x\|_2$ geometrically represents the length of the vector x

Matrix norms

Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$

Definition (Induced norm)

We define by

$$\|A\| = \max_{\substack{x\neq 0 \\ x\in\mathbb{C}^n}} \frac{\|Ax\|}{\|x\|}$$

the matrix norm of A *induced* by a vector norm $\|\cdot\|$ in $\mathbb{C}^n$.

NTNU
Norwegian University of
Science and Technology

Matrix norms

Examples

$$\|A\|_1 = \max_{\substack{x \neq 0 \\ x \in \mathbb{C}^n}} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}| \qquad \text{(induced 1-nom)}$$

$$\|A\|_\infty = \max_{\substack{x \neq 0 \\ x \in \mathbb{C}^n}} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| \qquad \text{(induced } \infty\text{-norm)}$$

$$\|A\|_2 = \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)} \qquad \text{(induced 2-norm)}$$

where $\rho()$ denotes the spectral radius.

NTNU
Norwegian University of
Science and Technology

Matrix norms

Frobenius norm

$$\|A\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2} = \sqrt{tr(A^*A)} = \sqrt{tr(AA^*)}$$

where $tr()$ represents the trace

Inequalities

- $\|A\| \geq 0, \quad \forall A \in \mathbb{C}^{n \times n}$ and $\|A\| = 0 \iff A = 0$
- $\|\alpha A\| = |\alpha| \, \|A\|, \quad \forall A \in \mathbb{C}^{n \times n}, \alpha \in \mathbb{C}$
- $\|A + B\| \leq \|A\| + \|B\|, \quad \forall A, B \in \mathbb{C}^{n \times n}$
- $\|AB\| \leq \|A\| \, \|B\| \quad \forall A, B \in \mathbb{C}^{n \times n}$
- $\|Ax\| \leq \|A\| \, \|x\|, \quad \forall A \in \mathbb{C}^{n \times n}, \forall x \in \mathbb{C}^n$

NTNU
Norwegian University of
Science and Technology

Matrix-vector products

Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ and $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$.

Write

$$A = \left[a_1 \middle| a_2 \middle| \cdots \middle| a_n \right]$$

where $a_j \in \mathbb{C}^n$ represent the *columns* of A.
We have that

$$Ax = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n$$

i.e. Ax is a *linear combination of columns* of A with coefficient vector x.

NTNU
Norwegian University of
Science and Technology

Matrix-vector products

Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ and $\tilde{x} = [\tilde{x}_1, \ldots, \tilde{x}_n]$, a *row* vector.
Now write

$$A = \begin{bmatrix} \tilde{a}_1 \\ \hline \tilde{a}_2 \\ \hline \vdots \\ \hline \tilde{a}_n \end{bmatrix}$$

where $\tilde{a}_i$ represents row i.
Then

$$\tilde{x}A = \tilde{x}_1\tilde{a}_1 + \tilde{x}_2\tilde{a}_2 + \ldots + \tilde{x}_n\tilde{a}_n$$

i.e. $\tilde{x}A$ is a *linear combination of rows* of A with coefficient vector $\tilde{x}$.

NTNU
Norwegian University of
Science and Technology

Matrix-vector products

Matrix-matrix products

Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ and $B = (b_{ij}) \in \mathbb{C}^{n\times n}$, such that

$$A = \left[a_1 \middle| \cdots \middle| a_n \right] \quad \text{and} \quad B = \left[b_1 \middle| \cdots \middle| b_n \right].$$

We have

$$AB = \left[Ab_1 \middle| Ab_2 \middle| \cdots \middle| Ab_n \right]$$

i.e. the matrix-vector product Ab_j yields column j of the matrix-matrix product AB.

NTNU
Norwegian University of
Science and Technology

Some important classes of matrices

Assume $A \in \mathbb{C}^{n\times n}$

Hermitian (symmetric, for real A)

$$A^* = A \qquad \text{(self-adjoint)}$$

Positive-definite

$$x^*Ax > 0, \quad \forall x \in \mathbb{C},\ x \neq 0$$

Note: A is called *semi-positive-definite* if the inequality in not strict (i.e. $x^*Ax \geq 0 \quad \forall x \neq 0$).

NTNU
Norwegian University of Science and Technology

Some important classes of matrices

Assume $A \in \mathbb{C}^{n \times n}$

Unitary

$$A^*A = AA^* = I$$

The columns of A are are *orthogonal* with respect to the Euclidean inner product; i.e. $a_i^* a_j = 0$ for $i \neq j$. More precisely $a_i^* a_j = \delta_{ij}$ (*orthonormality*).

Q unitary matrix $\Longrightarrow \left\|Qx\right\|_2 = \|x\|_2 \quad \forall x \in \mathbb{C}^n$ (Q *preserves length*)

Also $\left\|Q\right\|_2 = 1$.

Normal

$A^*A = AA^*$

Observe that all unitary and Hermitian matrices are normal.

NTNU
Norwegian University of
Science and Technology

Some important classes of matrices

Assume $A \in \mathbb{C}^{n \times n}$

Regular or nonsingular

$$det(A) \neq 0$$

Sparse

Most of the entries are *zeros*.
Examples: Diagonal, tri-diagonal, banded etc

Idempotent

$$A^2 = A$$

Also called a *projector*

NTNU
Norwegian University of Science and Technology

SOME IMPORTANT CLASSES OF MATRICES

Assume $A \in \mathbb{C}^{n\times n}$

Triangular

Two types

1) Upper triangular: $a_{ij} = 0, \quad \forall i > j$
   i.e. all entries *below* the leading diagonal are zeros.

2) Lower triangular: $a_{ij} = 0, \quad \forall i < j$
   i.e. all entries *above* the leading diagonal are zeros.

EIGENVALUES

Assume $A \in \mathbb{C}^{n \times n}$

DEFINITION (EIGENVALUE & EIGENVECTOR)

An *eigenvector* of A refers to any *nonzero* vector $v \in \mathbb{C}^n$ such that

$$Av = \lambda v$$

for some scalar $\lambda \in \mathbb{C}$.
The corresponding scalar λ is called an *eigenvalue* of A.
Thus the pair (v, λ) is called an *eigenpair*.

Eigenvalues

Corresponding to each eigenvalue λ is a subspace

$$E_\lambda = \left\{ x \in \mathbb{C}^n \middle| Ax = \lambda x \right\}$$

called the *eigenspace*. It is the nullspace of the matrix $(\lambda I - A)$
E_λ is *invariant* under A (i.e. $AE_\lambda \subset E_\lambda$). Also $\dim E_\lambda \geq 1$.

We have the following definitions:

$\sigma(A) =$ the set of all eigenvalues of $A =$ *spectrum* of A

$\rho(A) = \max\limits_{\lambda \in \sigma(A)} |\lambda|$, is called the *spectral radius* of A.

$\kappa = \kappa(A) = \dfrac{\max\limits_{\lambda \in \sigma(A)} |\lambda|}{\min\limits_{\lambda \in \sigma(A)} |\lambda|}$, is called the *condition number* of A.

EIGENVALUES

Characteristic polynomial

$$p_A(z) := det(zI - A), \quad z \in \mathbb{C}$$

Roots of p_A form the eigenvalues of A;
i.e. λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$.

By the *Fundamental Theorem of Algebra* p_A has n roots in $\mathbb{C}$;
so we can write

$$p_A(z) = (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n)$$

Eigenvalues

Defective matrix

- The *algebraic multiplicity* of an eigenvalue λ refers to the number of times λ occurs as a root of p_A.
- The algebraic multiplicity is always *greater than or equal* to its *geometric multiplicity* (namely, the dimension of its eigenspace, i.e. $\dim E_\lambda$).
- A *simple* eigenvalue is one with algebraic multiplicity 1.

Definition (Defective matrix)

An eigenvalue whose algebraic multiplicity exceeds its geometric multiplicity is called *defective*.
A *defective matrix* is one that has a defective eigenvalue.

EIGENVALUES

Defective matrix

Example:

$$A = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix} \text{ is defective. } B = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix} \text{ is nondefective.}$$

All diagonal matrices are nondefective.
All diagonalizable matrices are nondefective, and vice-versa.

Question: What do we mean by *diagonalizable*?

Eigenvalues

Diagonalizable matrix

Definition (Diagonalizable matrix)

A matrix $A \in \mathbb{C}^{n\times n}$ is diagonalizable if there exists an invertible matrix X such that the product

$$X^{-1}AX = \Lambda$$

is a diagonal matrix.

Thus A and Λ are similar matrices.
The diagonal entries of $\Lambda = diag\left([\lambda_1, \dots, \lambda_n]\right)$ are the eigenvalues of A, while the columns of $X = \left[v_1 \middle| \dots \middle| v_n\right]$ are the corresponding eigenvectors.
Check that $AX = X\Lambda \Longrightarrow Av_j = \lambda_j v_j$.

NTNU
Norwegian University of Science and Technology

EIGENVALUES

THEOREM (GERSCHGORIN)

Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$*. For each* $i = 1,\dots,n$ *define* $r_i = \sum_{j=1}^{n} a_{ij}$*, and let* $C_i = \left\{ z \in \mathbb{C} \,\middle|\, |z - a_{ii}| \le r_i \right\}$ *represent a circular disc of radius* r_i *and center* a_{ii} *in the complex plane. Then*

1) *every eigenvalue of* A *lies in at least one of the discs* $C_i, \quad i = 1, \dots, n$*;*
2) *if the union of any* k *discs forms a connected domain* D_k *that is disjoint from the remaining* $(n - k)$ *discs, then there are precisely* k *eigenvalues within* D_k*.*

EIGENVALUES

Some important results

1) All the eigenvalues of a Hermitian matrix are real.
2) Suppose $\lambda_1, \ldots, \lambda_n$ denote all the eigenvalues of $A \in \mathbb{C}^{n\times n}$ (including multiplicities). Then

$$det(A) = \prod_{j=1}^{n} \lambda_j \quad \text{and} \quad tr(A) = \sum_{j=1}^{n} \lambda_j$$

   where $tr(A) = \sum_{i=1}^{n} a_{ii}$ is the *trace* of A.
3) All normal matrices A are *unitarily diagonalizable*. That is, there is exists a unitary matrix Q such that

$$Q^* A Q = \Lambda$$

   where Λ is diagonal.

NTNU
Norwegian University of Science and Technology

Eigenvalues

Some important results

4) In general $\rho(A) \leq \|A\|_2\,, \quad \forall A \in \mathbb{C}^{n\times n}$. However $\rho(A) = \|A\|_2$ if $A = A^*$.
5) For a unitary matrix Q,

$$|\lambda_Q| = 1, \quad \rho(Q) = 1, \quad \kappa(Q) = 1.$$

NTNU
Norwegian University of Science and Technology

Matrix decompositions

Schur decomposition (triangularization)

$$A = QRQ^*$$

where Q is unitary and R is upper triangular.
Type:[2] Any square matrix $A \in \mathbb{C}^{n\times n}$.

Eigenvalue decomposition (diagonalization)

$$A = X\Lambda X^{-1}$$

where $\Lambda = diag([\lambda_1, \dots, \lambda_n])$, $\quad X = \left[v_1 \middle| \cdots \middle| v_n\right]$ and (v_j, λ_j) are eigenpairs of A.
Type: $A \in \mathbb{C}^{n\times n}$ nondefective (e.g. all normal matrices).

[2] Type of matrices that can be factorized into this form

NTNU
Norwegian University of Science and Technology

Matrix decompositions

cholesky decomposition

$$A = LL^*$$

where L is lower triangular.

Type: $A \in \mathbb{C}^{n \times n}$ Hermitian and positive definite (PD).

LU decomposition

$$A = LU$$

where L is lower triangular and U is upper triangular.

Type: $A \in \mathbb{C}^{n \times n}$ but not all square matrices. However with *pivoting* all square matrices can be LU-factorized.

Algorithms for LU factorizations are simply modified versions of Gaussian elimination.

Matrix decompositions

QR decomposition

$$A = QR$$

where Q is unitary and R is upper triangular.

Type: $A \in \mathbb{C}^{n \times n}$, with possible extensions to rectangular matrices.
Some common algorithms for QR factorization include Gram-Schimdt, Householder reflections, Givens rotation

NTNU
Norwegian University of Science and Technology

Matrix decompositions

Singular value decomposition

$$A = U\Sigma V^*$$

where $U \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{m\times m}$ are unitary, $\Sigma \in \mathbb{R}^{n\times m}$ is diagonal.

Type: This is a more general diagonalization that applies to all matrices $A \in \mathbb{C}^{n\times m}$, even defective as well as rectangular matrices.
Σ is *uniquely* determined [Trefethen & Bau].

Suppose $A \in \mathbb{C}^{n\times m}$. We have

$$U = \left[u_1 \middle| \cdots \middle| u_n\right], \quad V = \left[v_1 \middle| \cdots \middle| v_m\right], \quad Au_j = \sigma_j v_j, \quad j = 1, \ldots, p$$

where $p = \min(n, m)$. The $\sigma_j \in \mathbb{R}$ are called the *singular values* of A.
The singular values form the diagonal entries of Σ. They are nonnegative and decreasing, i.e., $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$.

Matrix decompositions

Condition number

Given $A \in \mathbb{C}^{n \times n}$, A nonsingular, we define the condtion number κ as follows:

$$\kappa = \kappa(A) := \|A\| \, \|A^{-1}\|$$

where $\|\cdot\|$ denote a matrix norm. In relation to eigenvalues and singular values we get the more specifically that

1. $\kappa = \dfrac{\lambda_{\text{max}}}{\lambda_{\text{min}}}$, where $\lambda_{\text{max}} = \max|\lambda|$ and $\lambda_{\text{min}} = \min|\lambda|$.
2. $\kappa = \dfrac{\sigma_{\text{max}}}{\sigma_{\text{min}}}$, where $\sigma_{\text{max}} = \max\sigma$ and $\sigma_{\text{min}} = \min\sigma$.

Matrix decompositions

	Advantages	Type
Schur	•Obtain eigenvalues •efficient linear solvers	all square matrices
Diagonalization	•Obtain eigenvalues •efficient linear solvers •easy computation of matrix exponentials	nondefective matrices (e.g normal)
QR	•Obtain eigenvalues •efficient linear solvers	all matrices in $\mathbb{C}^{m\times n}$, $m \geq n$
Cholesky	efficient linear solvers	Hermitian & PD
LU (with pivoting)	efficient linear solvers	all square matrice
SVD	•Obtain singular values •efficient linear solvers •easy computation of matrix exponentials	all matrices in $\mathbb{C}^{m\times n}$

Similarity transforms

Let $X \in \mathbb{C}^{n\times n}$ be invertible. Any map of the form

$$A \longmapsto X^{-1}AX, \quad A \in \mathbb{C}^{n\times n}$$

is called a *similarity transform*.

Two matrices $A, B \in \mathbb{C}^{n\times n}$ are said to be *similar* if there exists and invertible matrix $X \in \mathbb{C}^{n\times n}$ such that $B = X^{-1}AX$.

Two matrices are similar if and only if they have the *same* eigenvalues and eigenvectors.

Observation: The Schur and eigenvalue decompositions are examples of similarity transforms.

Iterative methods

Given linear system

$$Ax = b$$

Make an *initial guess* for the solution x_0:
This introduces an error $e_0 = x - x_0$
We have

$$Ae_0 = b - Ax_0 = r_0$$

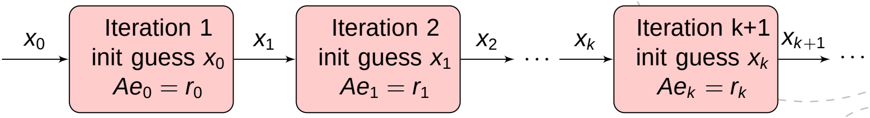
We call r_0 the *residual error.*
An iterative method uses x_0 to compute a *better* solution x_1 that has a smaller error and residual error. I.e.

$$||e_1|| \leq ||e_0||$$
$$||r_1|| \leq ||r_0||$$

where $e_k = x - x_k$ and $r_k = b - Ax_k = Ae_k, \quad k = 0, 1, \ldots.$

NTNU
Norwegian University of Science and Technology

Iterative methods



The method is *convergent* if we have that

$$\|e_k\| \leq c_k \|e_0\|, \qquad \text{for each } k = 0, 1, 2, \ldots$$

where $\{c_k\}$ is a nonnegative null sequence (i.e. $c_k \to 0$ as $k \to \infty$).

Iterative methods

- Iterative methods are more *effecient* for solving linear systems than direct methods.
- Direct methods involving Gaussian elimination or matrix decompositions spend *more computation time* and usually requires *more memory storage.*

Iterative methods

We shall study the following iterative methods

1) Basic methods:
   E.g. Jacobi, Gauss-Seidel, Successive overrelaxation
2) Projection methods:
   E.g. steepest descent, Minimum Residual (MRes)
3) Krylov subspace methods:
   - Generalized minimum residual (GMRes)
   - Lanczos/Anoldi algorithm
   - Conjugate gradient algorithm
4) Multigrid methods
5) Eigenvalue algorithms
6) QR factorization algorithms:
   E.g.Householder, Gram-Schmidt, Given etc