

# Numerical Methods of Linear Algebra for Sparse Matrices

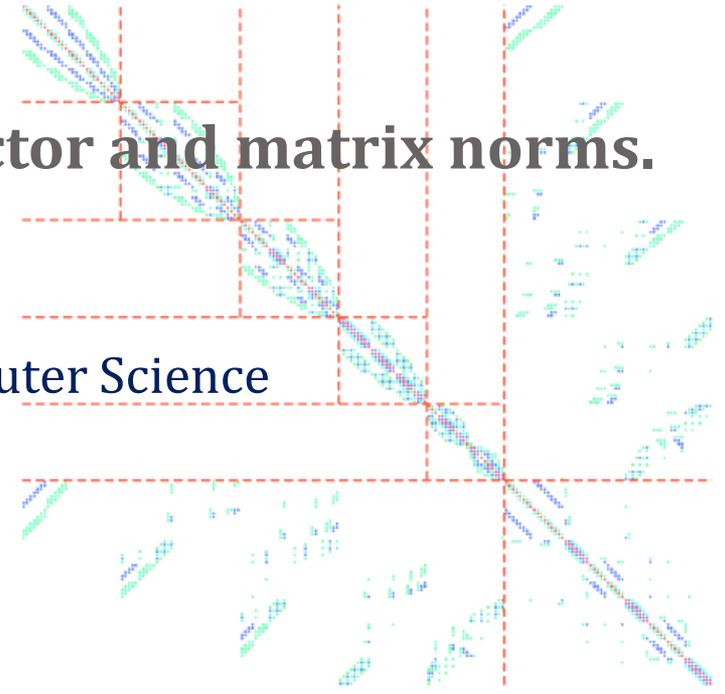
## Lecture 2

**Fundamentals of Linear Algebra. Vector and matrix norms.**  
**Anna Nasedkina, PhD, Assoc. prof.**

Department of Mathematical Modeling

Institute of Mathematics, Mechanics and Computer Science

Southern Federal University



# Outline

---

- Types of square matrices
- Structure of square matrices with respect to location of zero entries
- Vector and matrix norms

# Types of square matrices: symmetric and Hermitian

- The choice of solution method often depends on the structure of a matrix. For example, symmetry has an impact on the eigenstructure

- **Symmetric matrix**

$$A^T = A$$

- **Hermitian matrix**

$$A^H = A, \text{ where } A^H = \overline{A^T}$$

- **Skew-symmetric matrix**

$$A^T = -A$$

- **Skew-Hermitian matrix**

$$A^H = -A$$

## Properties of Hermitian matrices

- 1)  $\det A \in \mathbb{R}$  (determinant is a real scalar)
- 2) if  $A$  is Hermitian, then  $A^{-1}$  (if exists) is Hermitian
- 3) if  $A, B$  are Hermitian, then  $AB$  is Hermitian  $\Leftrightarrow AB = BA$   
 $\Leftrightarrow$  stands for "if and only if"
- 4) if  $A, B$  are Hermitian, then  $A + B$  is Hermitian

# Types of square matrices: positive and negative, normal, unitary and orthogonal

- **Positive matrix**  $a_{ij} > 0, i, j = \overline{1, n}$
- **Non-positive matrix**  $a_{ij} \leq 0, i, j = \overline{1, n}$
- **Negative matrix**  $a_{ij} < 0, i, j = \overline{1, n}$
- **Non-negative matrix**  $a_{ij} \geq 0, i, j = \overline{1, n}$

- **Normal matrix**

$$A^H A = A A^H \text{ (or } A^T A = A A^T \text{ for real matrices)}$$

- **Unitary matrix** (usually denoted as  $Q$ )

$$Q^H Q = Q Q^H = I, \text{ where } I \text{ is identity matrix}$$

**Property:**  $Q^{-1} = Q^H$

For nonsquare  $Q \in \mathbb{C}^{n \times m}$   
either  $Q^H Q = I_m$  (orthonormal columns)  
or  $Q Q^H = I_n$  (orthonormal rows)

- **Orthogonal matrix** is a unitary real matrix  $Q \in \mathbb{R}^{n \times n}$

$$Q^T Q = Q Q^T = I$$

**Property:**  $Q^{-1} = Q^T$

# Structures of square matrices with respect to location of zero entries: diagonal, triangular, Hessenberg

- **Diagonal matrix**  $a_{ij} = 0$  for  $\forall j \neq i$

$\forall \Leftrightarrow$  stands for "for every" or "for any"

Notation:  $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$

- **Upper-triangular matrix**

$$a_{ij} = 0, \text{ for } \forall i > j$$

|   |   |   |   |
|---|---|---|---|
| × | × | × | × |
| 0 | × | × | × |
| 0 | 0 | × | × |
| 0 | 0 | 0 | × |

- **Lower-triangular matrix**

$$a_{ij} = 0, \text{ for } \forall i < j$$

|   |   |   |   |
|---|---|---|---|
| × | 0 | 0 | 0 |
| × | × | 0 | 0 |
| × | × | × | 0 |
| × | × | × | × |

- **Upper-Hessenberg matrix**

$$a_{ij} = 0, \text{ for } \forall i, j: i > j+1$$

|   |   |   |   |
|---|---|---|---|
| × | × | × | × |
| × | × | × | × |
| 0 | × | × | × |
| 0 | 0 | × | × |

- **Lower-Hessenberg matrix**

$$a_{ij} = 0, \text{ for } \forall i, j: i < j+1$$

|   |   |   |   |
|---|---|---|---|
| × | × | 0 | 0 |
| × | × | × | 0 |
| × | × | × | × |
| × | × | × | × |

# Structures of square matrices with respect to location of zero entries: bidiagonal, tridiagonal, banded

- **Upper-bidiagonal matrix**

$$a_{ij} = 0, \text{ for } \forall j \neq i \text{ or } j \neq i+1$$

- **Lower-bidiagonal matrix**

$$a_{ij} = 0, \text{ for } \forall j \neq i \text{ or } j \neq i-1$$

- **Tridiagonal matrix**

$$a_{ij} = 0, \text{ for } \forall i, j: |j-i| > 1$$

Notation:  $A = \text{tridiag}(a_{i,i-1}, a_{ii}, a_{i,i+1})$

- **Banded matrix**

$$a_{ij} \neq 0 \Leftrightarrow i - m_l \leq j \leq i + m_u, \text{ where } m_l, m_u \text{ are positive integers}$$

$m_l + m_u + 1$  is the bandwidth

$m_l$  is the lower bandwidth:  $a_{ij} = 0, \text{ for } \forall i > j + m_l$

$m_u$  is the upper bandwidth:  $a_{ij} = 0, \text{ for } \forall j > i + m_u$

# Structures of square matrices with respect to location of zero entries: block, block diagonal and tridiagonal

- **Block matrix**

$$A = \begin{pmatrix} A_{11} & \dots & A_{1l} \\ \dots & \dots & \dots \\ A_{k1} & \dots & A_{kl} \end{pmatrix}, \text{ where } k \times l \text{ are block sizes of } A$$

- **Block-diagonal matrix**

$$A_{ij} = 0, \text{ for } \forall j \neq i \text{ or } j \neq i - 1$$

- **Tridiagonal matrix**

$$A_{ij} = 0, \text{ for } \forall i, j: |j - i| > 1$$

$$\text{Notation: } A = \text{tridiag}(A_{i,i-1}, A_{ii}, A_{i,i+1})$$

- **Permutation matrix** is the identity matrix with its rows (or columns) permuted

Example

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

# Inner product

Let's consider a complex space  $X \subset \mathbb{C}^n$  ( $X$  is a subspace of  $\mathbb{C}^n$ ).

Note that  $\mathbb{C}^n$  is the space of vector-columns.

For the vectors  $x, y \in X$  an **inner product**

is a mapping  $s(x, y): X \times X \rightarrow \mathbb{C}$ , which satisfies the following properties:

1)  $s(x, y)$  is linear by  $x$

$$s(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 s(x_1, y) + \lambda_2 s(x_2, y) \quad \forall x_1, x_2 \in X, \lambda_1, \lambda_2 \in \mathbb{C}$$

2)  $s(x, y)$  is Hermitian

$$s(y, x) = \overline{s(x, y)} \quad \forall x, y \in X, \text{ therefore } s(x, x) \in \mathbb{R}$$

3)  $s(x, y)$  is positive definite:  $s(x, x) \geq 0$  and  $s(x, x) = 0 \Leftrightarrow x = 0$

$$s(x, 0) = 0 \quad \forall x \in X, \quad s(0, y) = 0 \quad \forall y \in X.$$

$$\text{Proof. } s(x, 0) = s(x, 0 \cdot y) = 0 \cdot s(x, y) = 0$$

**Cauchy - Schwartz inequality**

$$|s(x, y)|^2 \leq s(x, x) \cdot s(y, y)$$

# Euclidean inner product

When  $X = \mathbb{C}^n$ , the inner product  $s(x, y)$  is called **Euclidean** and denoted by  $(x, y)$ .

For the vector-columns  $x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \in \mathbb{C}^n$ ,  $y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \in \mathbb{C}^n$

$$(x, y) = \sum_{i=1}^n x_i \bar{y}_i = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n$$

In a matrix form (using matrix multiplication):  $(x, y) = y^H x$  ( $y^H$  is vector-row)

For  $X = \mathbb{R}^n$   $(x, y) = \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$ , in a matrix form :  $(x, y) = y^T x$

$$(x, x) = \sum_{i=1}^n x_i^2 = x_1^2 + x_2^2 + \dots + x_n^2$$

## Property

$$(Ax, y) = (x, A^H y) \geq 0 \quad \forall x \in \mathbb{C}^n, y \in \mathbb{C}^m, A \in \mathbb{C}^{m \times n}$$

**Proposition.** Unitary matrices preserve inner product:

$$(Qx, Qy) = (x, y) \quad \forall Q \in \mathbb{C}^{n \times n}$$

$$\triangle (Qx, Qy) = (x, Q^H Qy) = (x, Iy) = (x, y) \quad \square$$

$\triangle$  is the beginning of proof

$\square$  is the end of proof

# Vector norms

- Norms are needed to measure lengths of vectors and closeness of two vectors (or matrices).
- Examples of using vector norms: estimate convergence rate of an iterative method; estimate the error of an approximation to a given solution

**Vector norm** on  $X \subset \mathbb{C}^n$  is a function  $X \rightarrow \mathbb{R}$ , denoted by  $\|x\|$ ,

which satisfies the following properties:

1)  $\|x\| \geq 0 \quad \forall x \in X$  (the norm is non-negative)

$$\|x\| = 0 \Leftrightarrow x = 0$$

2)  $\|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall x \in X, \alpha \in \mathbb{C}$

3) triangle inequality

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in X$$

# Particular cases of vector norms

When  $X = \mathbb{C}^n$ , the **Euclidian norm** of  $x \in X$  is denoted by  $\|x\|_2$  and defined by

$$\|x\|_2 = \sqrt{(x, x)}$$

$$\text{Hence } (x, x) = \|x\|_2^2$$

**Proposition.** Unitary matrices preserve the Euclidian norm:

$$\|Qx\|_2 = \|x\|_2 \quad \forall Q \in \mathbb{C}^{n \times n}$$

- **Hölder norms** are most commonly used

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}$$

$$1) p = 1 \Rightarrow \|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$$

$$2) p = 2 \Rightarrow \text{Euclidian norm } \|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$$

$$3) p = \infty \Rightarrow \text{infinity norm } \|x\|_\infty = \max_{i=1, n} |x_i|$$

**Cauchy - Schwartz inequality**  
for Euclidian inner product and norm

$$|(x, y)| \leq \|x\|_2 \cdot \|y\|_2$$

# Matrix norms

- Matrix norms can be defined by vector norms, if we represent  $n$  by  $m$  matrices as vectors of the size  $nm$

For matrix  $A \in \mathbb{C}^{n \times m}$   $\|A\|$  is the **matrix norm**, if it satisfies the following properties:

1)  $\|A\| \geq 0 \quad \forall A \in \mathbb{C}^{n \times m}$  (the norm is non-negative)

$$\|A\| = 0 \Leftrightarrow A = 0$$

2)  $\|\alpha A\| = |\alpha| \cdot \|A\| \quad \forall A \in \mathbb{C}^{n \times m}, \alpha \in \mathbb{C}$

3) triangle inequality

$$\|A + B\| \leq \|A\| + \|B\| \quad \forall A, B \in \mathbb{C}^{n \times m}$$

For matrices, which admit matrix product, there is extra property:

4) multiplicative property

$$\|AB\| \leq \|A\| \cdot \|B\|$$

# Operator matrix norms

- Often the term “matrix norm” refers to an operator matrix norm. These operator norms are considered as proper matrix norms.

The set of matrix norms  $\|\cdot\|_{pq}$  are called **operator matrix norms**, if they are induced by two vector norms  $\|\cdot\|_p$  and  $\|\cdot\|_q$

$$\|A\|_{pq} = \max_{x \in \mathbb{C}^m, x \neq 0} \frac{\|Ax\|_p}{\|x\|_q} \quad \text{for } A \in \mathbb{C}^{n \times m}$$

When  $p = q$  we have  $\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p$

- **Consistency of norms**

Matrix p-norm is consistent with a vector p-norm, if

$$\|Ax\|_p \leq \|A\|_p \cdot \|x\|_p$$

Property.  $\forall A \in \mathbb{C}^{n \times m} \quad \|A^k\|_p = \|A\|_p^k$

# Particular cases of non-operator matrix norms

- Hölder matrix norms

$$\|x\|_{l_p} = \left( \sum_{j=1}^m \sum_{i=1}^n |a_{ij}|^p \right)^{1/p}, \quad \forall A \in \mathbb{C}^{n \times m}, p \geq 1$$

1)  $p = 1 \Rightarrow l_1$ -norm  $\|A\|_{l_1} = \sum_{j=1}^m \sum_{i=1}^n |a_{ij}| \quad \forall A \in \mathbb{C}^{n \times m}$

2)  $p = 2 \Rightarrow$  Frobenius norm  $\|A\|_F = \sqrt{\sum_{j=1}^m \sum_{i=1}^n |a_{ij}|^2}$

Also  $\|A\|_F = \sqrt{\text{tr}(A^H A)} = \sqrt{\text{tr}(A A^H)}$

Property. Frobenius norm is consistent with the vector norm  $\|\cdot\|_2$

For square matrices  $A \in \mathbb{C}^{n \times n}$  this norm is called Euclidian norm

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

- **M-norms**  $\|A\|_M = n \cdot \max_{i,j=1,n} |a_{ij}|$  for square matrices  $A \in \mathbb{C}^{n \times m}$

# Particular cases of operator matrix norms

## 1) maximal column norm

$$\|A\|_1 = \max_{j=1,m} \sum_{i=1}^n |a_{ij}| \quad \forall A \in \mathbb{C}^{n \times m}$$

## 2) maximal row norm

$$\|A\|_\infty = \max_{i=1,n} \sum_{j=1}^m |a_{ij}| \quad \forall A \in \mathbb{C}^{n \times m}$$

## 3) spectral norm

$$\|A\|_2 = \sqrt{\rho(AA^H)} = \sqrt{\rho(A^H A)} \quad \forall A \in \mathbb{C}^{n \times m}$$

where  $\rho$  is the spectral radius (maximal eigenvalue by modulus)

It is known that all eigenvalues (the spectrum) of  $A^H A$

are nonnegative:  $\sigma(A^H A) \geq 0$

Also  $\|A\|_2 = \max_{i=1,n} \sigma_i$  where  $\sigma_i$  is the **singular value**,

i.e.  $\sigma_i^2$  is an eigenvalue of  $A^H A$

## Important relations

- 1) Matrix spectral norm  $\|\cdot\|_2$  is consistent with vector norm  $\|\cdot\|_2$ ,  
maximal column norm  $\|\cdot\|_1$  is consistent with vector norm  $\|\cdot\|_1$ ,  
maximal row norm  $\|\cdot\|_\infty$  is consistent with vector norm  $\|\cdot\|_\infty$
- 2) Matrix M-norm  $\|\cdot\|_2$  is consistent with vector norms  $\|\cdot\|_1$ ,  $\|\cdot\|_2$ ,  $\|\cdot\|_\infty$
- 3) Multiplicative property  $\|AB\| \leq \|A\| \cdot \|B\|$  holds for any operator norm  
and for Frobenius norm.
- 4)  $\rho(A) \leq \|A\|$  for any matrix norm (the spectral radius is not greater  
than the norm).

 Compute the  $p$ -norm for  $p = 1, 2, \infty, F$  for the matrix

$$A = \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}$$