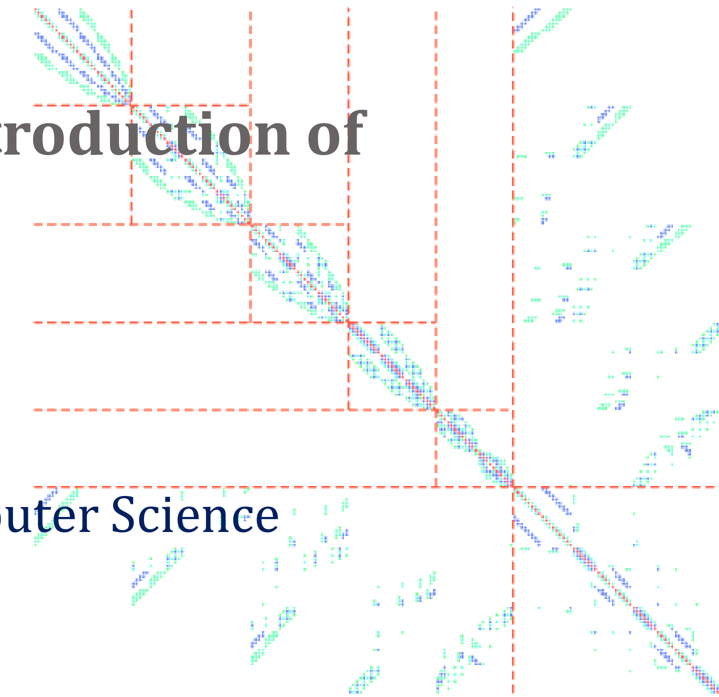
Numerical Methods of Linear Algebra for Sparse Matrices

Lecture 1

Fundamentals of Linear Algebra. Introduction of elementary notation



Anna Nasedkina, PhD, Assoc. prof.
Department of Mathematical Modeling
Institute of Mathematics, Mechanics and Computer Science
Southern Federal University

Outline

- Matrices and matrix algorithms in modern world
- Linear systems and sparse linear systems
- Fundamentals of Linear Algebra (review of basic matrix theory, introduction of elementary notation)

Matrices and matrix algorithms in modern world

- Matrix algorithms pervade most areas of science and engineering
- Matrix algorithms are used by most disciplines requiring numerical computing
- Computations involving matrices is the topic becoming increasingly important in Computer Science
- Matrix algorithms in computer science are used for data mining, information retrieval, search engines, pattern recognition, graphics, ...

Example of a linear system

- **Problem** (from Yosef Saad lecture). A set of 12 coins containing nickels (5c each), dimes (10c each) and quarters (25c each) totals to \$1.45. In addition, the total without the nickels amounts to \$1.25. How many of each coin are there?
- **Solution.** Let x_n be the number of nickels, x_d be the number of dimes, x_q be the number of quarters

- **Equation form**

1st equation: $x_n + x_d + x_q = 12$

2nd equation: $5x_n + 10x_d + 25x_q = 145$

3rd equation: $0x_n + 10x_d + 25x_q = 125$

What's the answer?

- **Matrix form**

$$\begin{pmatrix} 1 & 1 & 1 \\ 5 & 10 & 25 \\ 0 & 10 & 25 \end{pmatrix} \begin{pmatrix} x_n \\ x_d \\ x_q \end{pmatrix} = \begin{pmatrix} 12 \\ 145 \\ 125 \end{pmatrix}$$

Matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 10 & 25 \\ 0 & 10 & 25 \end{pmatrix}$

Right-hand side vector $b = \begin{pmatrix} 12 \\ 145 \\ 125 \end{pmatrix}$

Vector of unknowns $x = \begin{pmatrix} x_n \\ x_d \\ x_q \end{pmatrix}$

Matrix equation: $Ax = b$

Solve in Matlab

- Matrix form:

$$\begin{pmatrix} 1 & 1 & 1 \\ 5 & 10 & 25 \\ 0 & 10 & 25 \end{pmatrix} \begin{pmatrix} x_n \\ x_d \\ x_q \end{pmatrix} = \begin{pmatrix} 12 \\ 145 \\ 125 \end{pmatrix}$$

- **Answer**: 4 nickels, 5 dimes and 3 quarters

- Question: is the order of equations important or not?
- Will the solution change?

```
A=[1 1 1; 5 10 25; 0 10 25]
```

```
A = 3×3
     1     1     1
     5    10    25
     0    10    25
```

```
b=[12; 145; 125]
```

```
b = 3×1
    12
   145
   125
```

```
x=A\b
```

```
x = 3×1
     4
     5
     3
```

Experiment in Matlab

```
A=[1 1 1; 5 10 25; 0 10 25]
b=[12; 145; 125]
x=A\b
```

```
p=[2 1 3]
A=A(p,:)
b=b(p,:)
x=A\b
```

- **Question**: is the order of equations important or not? Will the solution change?
- **Answer**: No!

```
x = 3×1
     4
     5
     3

p = 1×3
     2     1     3

A = 3×3
     5    10    25
     1     1     1
     0    10    25

b = 3×1
   145
    12
   125

x = 3×1
     4
     5
     3
```

Source of large linear systems

- Typical Large-scale problem (ex. Fluid flow)

Physical Model

↓

Mathematical Model:

Nonlinear Partial Differential Equations (PDEs)

↓

Discretization by Numerical Method:

Finite difference method (FDM), Finite element method (FEM), etc.

↓

Linearization (ex. Newton-Raphson)

↓

Sequence of Sparse Linear Systems $\boldsymbol{Ax = b}$

↓

Discrete solution (find x)

What is a PDE?

- A Partial Differential Equation (PDE) is a relationship between an unknown function of several variables and its partial derivatives.
- $u = u(x_1, x_2, x_3, t)$ is unknown function (field variable), x_1, x_2, x_3 (space coordinates) and t (time) are independent variables

Order of a PDE

the highest derivative in the equation

$$\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0 \quad \text{is a first-order PDE.}$$

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = 0 \quad \text{is a second-order PDE.}$$

$$\frac{\partial^4 u}{\partial x_1^4} + \frac{\partial^2 u}{\partial x_2^2} - u = 0 \quad \text{is a fourth-order PDE.}$$

$$\left(\frac{\partial u}{\partial x_1}\right)^3 + \frac{\partial u}{\partial x_2} + u^4 = 0 \quad \text{is a first-order PDE.}$$

Linear and nonlinear PDEs

A *linear* PDE is of the first degree both in its field variables and partial derivatives

$$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} = 0 \quad \text{is linear .}$$

$$\frac{\partial u}{\partial x_1} + \left(\frac{\partial u}{\partial x_2}\right)^2 = 0 \quad \text{is nonlinear .}$$

$$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} + u^2 = 0 \quad \text{is nonlinear .}$$

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = x_1 \quad \text{is linear .}$$

$$\frac{\partial^2 u}{\partial x_1^2} + u\frac{\partial^2 u}{\partial x_2^2} = 0 \quad \text{is quasilinear .}$$

https://en.wikiversity.org/wiki/Nonlinear_finite_elements/Partial_differential_equations

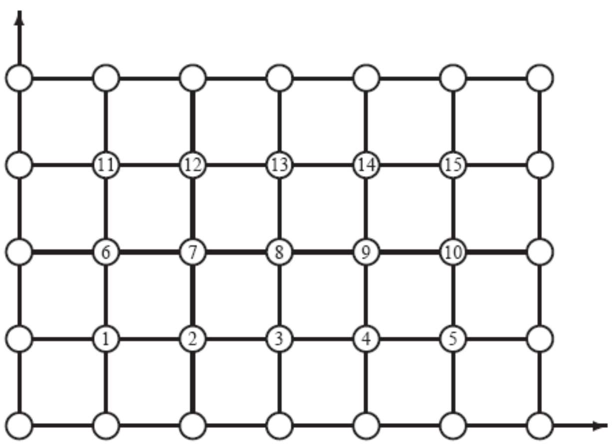
What is a sparse matrix?

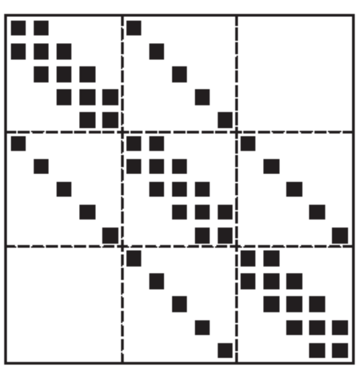
- A sparse matrix is a matrix which has very few nonzero elements.
- Sparsity is the fraction of zero elements in a matrix
- Density is the fraction of non-zero elements in a matrix
- **Example of sparse matrix**: 64 elements, 52 zero elements and 12 nonzero elements (82% sparsity, 18% density)

$$\begin{pmatrix} 1.0 & 0 & 5.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.0 & 0 & 0 & 0 & 0 & 11.0 & 0 \\ 0 & 0 & 0 & 0 & 9.0 & 0 & 0 & 0 \\ 0 & 0 & 6.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7.0 & 0 & 0 & 0 & 0 \\ 2.0 & 0 & 0 & 0 & 0 & 10.0 & 0 & 0 \\ 0 & 0 & 0 & 8.0 & 0 & 0 & 0 & 0 \\ 0 & 4.0 & 0 & 0 & 0 & 0 & 0 & 12.0 \end{pmatrix}$$

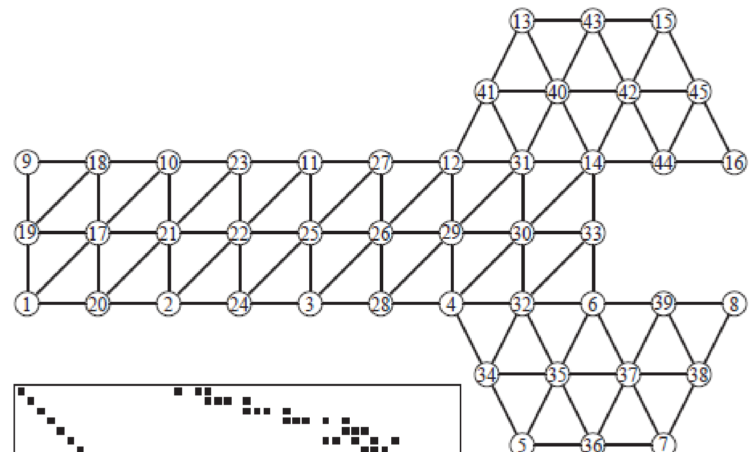
Discretization of PDEs

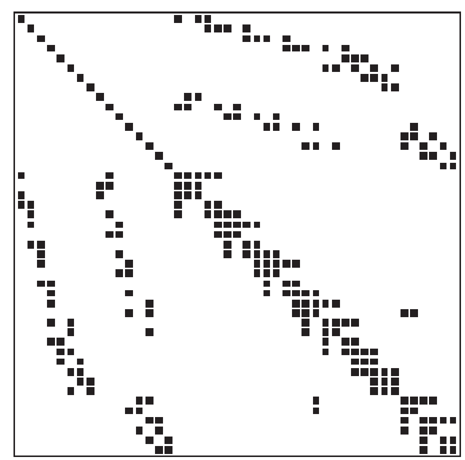
- Example: grid in finite difference method and corresponding sparse matrix
- Result: *structured* sparse matrix (nonzero entries form a regular pattern). Here we see blocks of diagonals





- Example: finite element mesh and corresponding sparse matrix
- Result: *unstructured* sparse matrix (nonzero entries are located irregularly). Although, we can spot a symmetric pattern here!





Definition of a matrix

A complex matrix $A = a_{ij}, \; A \in \mathbb{C}^{n \times m}$ is an array $n \times m$ of complex numbers,

A real matrix $A = a_{ij}, \; A \in \mathbb{R}^{n \times m}$ is an array $n \times m$ of real numbers,

where n is the number of rows, m is the number of columns

i is the row index, $i = 1, 2, \ldots, n$

j is the column index, $j = 1, 2, \ldots, m$

a_{ij} are the entries of the matrix

$\in$ stands for "belongs to"

Row-vector of a matrix

$a_{i\,*} = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ is the i-th row

Column-vector of a matrix

$$a_{*\,j} = \begin{pmatrix} a_{1\,j} \\ a_{2\,j} \\ \vdots \\ a_{n\,j} \end{pmatrix}$$ is the j-th column

$$A = \begin{pmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \end{pmatrix}$$

rows: 1, 2

columns: 1, 2, 3

Matrix A has

$n = 2$ rows, $m = 3$ columns

Matrix coefficients (entries) are

a_{11}=10, a_{12}=20, a_{13}=30,

a_{21}=40, a_{22}=50, a_{23}=60

Matrices in Matlab

```
% , or space separates columns
% ; separates rows

% Real matrices
A=[10,20,30;40,50,60]
B=[1 2;3 4]


a12=A(1,2) % matrix element

A_row2=A(2,:) % 2nd row

A_col3=A(:,3) % 3rd column



B_row_first=B(1,:) % 1st row

B_col_last=B(:,end) % last column



%complex matrix
A=[1+2i 3-1i; 4+5i -1i]
```

```
A = 2x3
    10    20    30
    40    50    60

B = 2x2
     1     2
     3     4

a12 = 20
A_row2 = 1x3
    40    50    60

A_col3 = 2x1
    30
    60

B_row_first = 1x2
     1     2

B_col_last = 2x1
     2
     4

A = 2x2 complex
   1.0000 + 2.0000i   3.0000 - 1.0000i
   4.0000 + 5.0000i   0.0000 - 1.0000i
```

Main operations with matrices

Addition

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

$$c_{ij} = a_{ij} + b_{ij}, \quad i = \overline{1, n}, \quad j = \overline{1, m}$$

Multiplication by a scalar

$$C = \alpha A, \quad A, C \in \mathbb{C}^{n\times m}, \quad \alpha \in \mathbb{C}$$

$$c_{ij} = \alpha a_{ij}, \quad i = \overline{1, n}, \quad j = \overline{1, m}$$

Multiplication by a matrix

$$C = AB, \quad A \in \mathbb{C}^{n\times m}, \quad B \in \mathbb{C}^{m\times p}, \quad C \in \mathbb{C}^{n\times p}$$

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}, \quad i = \overline{1, n}, \quad j = \overline{1, p}$$

Matrix multiplication

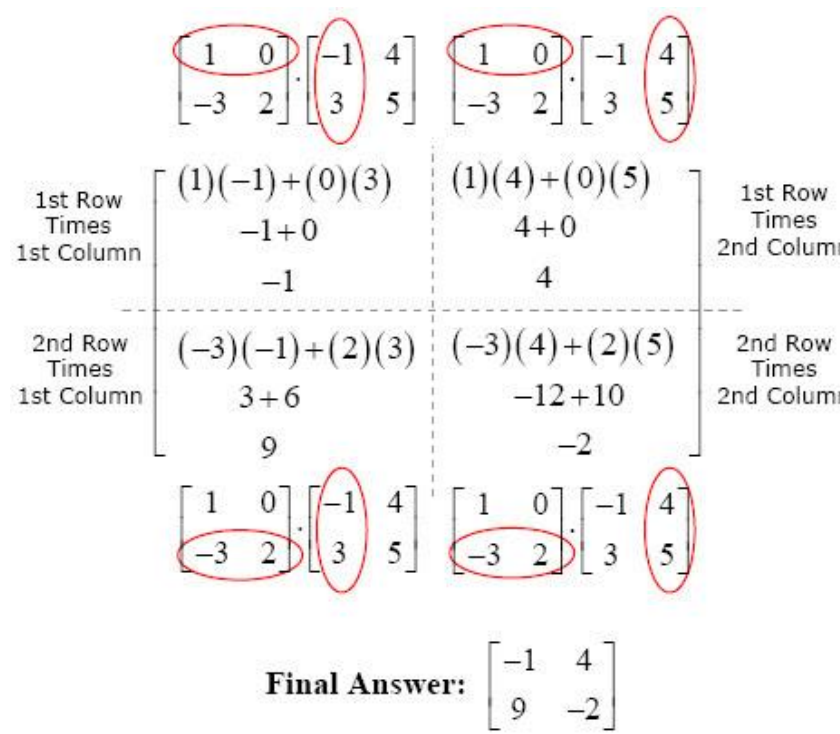
Multiplication by a matrix

$$C = AB, \quad A \in \mathbb{C}^{n\times m}, \quad B \in \mathbb{C}^{m\times p}, \quad C \in \mathbb{C}^{n\times p}$$

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}, \quad i = \overline{1,n}, \quad j = \overline{1,p}$$

True or false?

- $AB = BA$
- $A(BC) = (AB)C$



http://www.gradeamathhelp.com/matrix-multiplication.html

Is it possible to multiply A by B in these cases?

1) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} -5 & 6 & 7.5 \\ 0 & 3.1 & 2 \end{pmatrix}$

2) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 10 & 20 \\ 30 & 40 \end{pmatrix}$

Matrix operations in Matlab

```
A=[10,20,30; 40,50,60]
B=[1 2 3; 4 5 6]

% addition
C_ad=A+B

% multipication by a scalar
C_ms=3.6*A

% random matrices
A1=rand(2,3)
B1=rand(3,2)

% matrix multiplication
C1=A1*B1

C=A*B
```

```
A = 2x3
    10    20    30
    40    50    60

B = 2x3
     1     2     3
     4     5     6

C_ad = 2x3
    11    22    33
    44    55    66

C_ms = 2x3
    36    72   108
   144   180   216

A1 = 2x3
    0.7513    0.5060    0.8909
    0.2551    0.6991    0.9593

B1 = 3x2
    0.5472    0.2575
    0.1386    0.8407
    0.1493    0.2543

C1 = 2x2
    0.6142    0.8454
    0.3797    0.8973

Error using  *
Incorrect dimensions for matrix multiplication.
```

Matrix-matrix and matrix-vector product

Matrix product $C = AB,\ A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times p},\ C \in \mathbb{C}^{m\times p}$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},\quad i = \overline{1,m},\quad j = \overline{1,p}$$

- Column-wise product

$$c_{*j} = \sum_{k=1}^{n} b_{kj} a_{*k},\ c_{*j} = \begin{pmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{mj} \end{pmatrix} \text{ is the } j\text{-th column, } a_{*k} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} \text{ is the } k\text{-th column}$$

- Row-wise product

$$c_{i*} = \sum_{k=1}^{n} a_{ik} b_{k*},\ c_{i*} = \begin{pmatrix} c_{i1} & c_{i2} & \cdots & c_{ip} \end{pmatrix} \text{ is the } i\text{-th row,}$$

$$b_{k*} = \begin{pmatrix} b_{k1} & b_{k2} & \cdots & b_{kp} \end{pmatrix} \text{ is the } k\text{-th row}$$

What happens to these 3 different approaches when B has one column ($p = 1$)?

Characterize the matrices AA^T and A^TA when A is of dimension $n \times 1$.

- Sum of "outer-product" matrices

$$C = \sum_{k=1}^{n} a_{*k} b_{k*},\ a_{*k} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} \text{ is the } k\text{-th column, } b_{k*} = \begin{pmatrix} b_{k1} & b_{k2} & \cdots & b_{kp} \end{pmatrix} \text{ is the } k\text{-th row}$$

Transpose of a matrix

- Transpose of a matrix

For $A \in \mathbb{C}^{n\times m}$ its transpose is a matrix $A^T = C \in \mathbb{C}^{m\times n}$,

where $c_{ij} = a_{ji}, \ i = \overline{1,m}; \ j = \overline{1,n}$

Properties

1) $(AB)^T = B^T A^T; \quad (AB)^H = B^H A^H$

2) $(\alpha A)^T = \alpha A^T; \quad (\alpha A)^H = \overline{\alpha} A^H$

3) $(A^T)^T = A; \quad (A^H)^H = A$

4) $(A+B)^T = A^T + B^T$

- Transpose complex conjugate is

$A^H = \overline{A}^T = \overline{A^T}$

The bar denotes element-wise complex conjugation.

$(ABC)^T = ?$

True or false?

- $AA^T = A^T A$

Given a complex number	Its conjugate
$a+bi$	$a-bi$
$a-bi$	$a+bi$

$z = 2+3i \ \Rightarrow \ \overline{z} = 2-3i$

Transpose in Matlab

```
% Random integer matrix
A=randi(10,2,3)

% Transposition for real matrix
A'


% Random complex matrix
A=rand(2)+1i*rand(2)


%Transpose
At=transpose(A)



%Complex conjugate transpose
Act=ctranspose(A)
Act=A'
```

```
A = 2x3
     4     8    10
     7     1     8

ans = 3x2
     4     7
     8     1
    10     8

A = 2x2 complex
   0.4868 + 0.5085i   0.4468 + 0.8176i
   0.4359 + 0.5108i   0.3063 + 0.7948i

At = 2x2 complex
   0.4868 + 0.5085i   0.4359 + 0.5108i
   0.4468 + 0.8176i   0.3063 + 0.7948i

Act = 2x2 complex
   0.4868 - 0.5085i   0.4359 - 0.5108i
   0.4468 - 0.8176i   0.3063 - 0.7948i

Act = 2x2 complex
   0.4868 - 0.5085i   0.4359 - 0.5108i
   0.4468 - 0.8176i   0.3063 - 0.7948i
```

Square matrix, identity matrix and inverse

- The matrix is called square, if it has the same number of rows

and columns: $n = m$, so $A \in \mathbb{C}^{n \times n}$ or $A \in \mathbb{R}^{n \times n}$

- Identity matrix is the square* matrix $I = \{\delta_{ij}\},\ i, j = \overline{1, n}$ with

the elements of Kroneker delta: $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Property of identity matrix: $AI = IA = A \ \forall A \in \mathbb{C}^{n \times n}$

$\forall$ stands for "for any"

- Inverse of a matrix $A \in \mathbb{C}^{n \times n}$ (if exists) is the matrix $A^{-1} = C \in \mathbb{C}^{n \times n}$,

such that $CA = AC = I$

Properties of inverse

1) $(AB)^{-1} = B^{-1} A^{-1}$

2) $(\alpha A)^{-1} = \dfrac{1}{\alpha} A^{-1}$

3) $(A^T)^{-1} = (A^{-1})^T$

*Note. In some cases we can define rectangular identity matrix $I \in \mathbb{R}^{n \times m}$ with additional zero rows, if $n > m$ or additional zero columns, if $n < m$

Determinant of a matrix, singular matrix

- Recursive definition of a determinant for matrix $A \in \mathbb{C}^{n\times n}$

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j}), \text{ where } A_{1j} \in \mathbb{C}^{(n-1)\times(n-1)} \text{ is a matrix}$$

obtained by deletion of the 1st row and j-th column

Determinant of transpose and inverse

$$\det(A^{-1}) = \frac{1}{\det(A)}; \quad \det(A^T) = \det(A)$$

Determinant of identity matrix $\det(I) = 1$

Properties of determinant

1) $\det(AB) = \det(A)\det(B)$

2) $\det(\alpha A) = \alpha^n \det(A), \; \alpha \in \mathbb{C}$

3) $\det(\overline{A}) = \overline{\det(A)}$

- singular matrix: $\det(A) = 0$, the inverse of such matrix does not exist

Otherwise, when $\det(A) \neq 0$, the matrix is called nonsingular and $\exists\, A^{-1}$

Eigenvalues and eigenvectors

- A complex scalar λ is called an eigenvalue of a square matrix $A \in \mathbb{C}^{n\times n}$,

if $\exists\, u \in \mathbb{C}^n,\ u \neq 0:\ Au = \lambda u$

$\exists$ stands for "exists"

- The vector u is called eigenvector of the matrix A, associated with λ
- The set of all eigenvalues is the spectrum of A: $\sigma(A) = \{\lambda_1, \lambda_2, ..., \lambda_n\}$

The matrix of the size n has maximum n different eigenvalues

- Characteristic polynomial of the matrix $A \in \mathbb{C}^{n\times n}$

$p_A(\lambda) = \det(A - \lambda I)$ is the polynomial of degree n

Proposition 1. $\lambda \in \sigma(A)$, i.e. λ is an eigenvalue of A $\Leftrightarrow$ $\det(A - \lambda I) = 0$

$\Leftrightarrow$ stands for "if and only if"

Proposition 2. $\lambda \in \sigma(A) \Leftrightarrow \overline{\lambda} \in \sigma(A^H)$

Proposition 3. $\lambda \in \sigma(A) \Leftrightarrow \lambda$ is the root of the characteristic polynomial

Spectral radius and trace

- Spectral radius of a square matrix $A \in \mathbb{C}^{n\times n}$ is the maximum modulus of all eigenvalues:

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|$$

- Trace of the matrix $A \in \mathbb{C}^{n\times n}$ is the sum of all diagonal elements

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

Trace $tr(A)$ is equal to the *sum* of eigenvalues counted with their multiplicities as the roots of characteristic polynomial $p_A(\lambda)$.

Determinant $\det(A)$ is equal to the *product* of eigenvalues counted with their multiplicities as the roots of characteristic polynomial $p_A(\lambda)$.

✍ Trace, spectral radius, and determinant of

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}.$$