1)

Lecture 3

Canonical forms by similarity transformation

If $A \in \mathbb{C}^{n \times n}$ is a square matrix, sometimes we can reduce it to a simpler form (diagonal, bidiagonal, triangular, etc.)
Reduction means a similarity transformation, which preserves eigenvalues.

⑤ Similarity transformation

преобразование подобия

mapping $B \to A$: $\quad A = X B X^{-1}$

X is nonsingular

$\sigma(A) = \sigma(B)$ eigenvalues are the same

u_B is eigenvector of B $\Rightarrow u_A = X u_B$ eigenvector of A

u_A is eigenvector of $A \Rightarrow u_B = X^{-1} u_A$ is eigenvector of B

Matrices A and B are called similar

⑥ Reduction to a diagonal form

$$A = X \mathcal{D} X^{-1}$$

$$\mathcal{D} = \mathrm{diag}(\lambda_1, \dots, \lambda_n), \; \lambda_i \in \sigma(A)$$

When such reduction is possible?
Let's first give some definitions of eigenvalue multiplicities (кратность)

Multiplicities of eigenvalues

Ⓓ algebraic multiplicity of $\lambda \in \sigma(A)$ is a multiplicity of λ as a root of

$p_A(\lambda) = \det(A - \lambda I)$

$\mu(\lambda)$ denotes algebraic multiplicity

Ⓓ geometric multiplicity of $\lambda \in \sigma(A)$ is a number of lin. ind. eigenvectors:

$$\nu(\lambda) \overset{def}{=} \dim(\underbrace{Null(A - \lambda I)}_{\text{eigenspace}})$$

Ⓓ simple λ : $\mu(\lambda) = 1$

Ⓓ multiple λ : $\mu(\lambda) > 1$

Ⓓ semisimple λ : $\nu(\lambda) = \mu(\lambda)$

Ⓓ defective λ : $\nu(\lambda) < \mu(\lambda)$ ⇐ λ is multiple

Ⓓ deragotary λ : $\nu(\lambda) > 1$

Ex. If $\mu(\lambda) = 3$, $\nu(\lambda) = 2$, then λ is defective and deragotory.

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

$\mu(\lambda) = 2$, $\nu(\lambda) = 1$ $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

③ Ⓓ <u>simple matrix</u>: $\forall \lambda \in \sigma(A)$ is simple

Ⓓ <u>semisimple matrix</u> : $\forall \lambda \in \sigma(A)$ is semisimple

Ⓓ <u>derogatory matrix</u> (неполная)

$\exists \lambda : \nu(\lambda) > 1$ ($\exists$ derogatory λ)

Ⓓ <u>defective matrix</u>

$\exists \lambda : \nu(\lambda) < \mu(\lambda)$ ($\exists$ defective λ)

Note that $\forall \lambda \quad \nu(\lambda) \leq \mu(\lambda)$

Ⓣ || For similar matrices A and B $\sigma(A) = \sigma(B)$, hence multiplicities are the same for corresponding eigenvalues

Ⓣ || For a simple matrix all eigenvalues λ_i are distinct, hence all eigenvectors u_i are linear independent

---

<u>Canonical forms and factorizations</u>

1) Diagonal
2) Jordan
3) Schur
4) SVD (not a similarity transformation)
5) LU
6) Cholesky

---

1) <u>Diagonal form</u>

Ⓓ Matrix $A \in \mathbb{C}^{n \times n}$ is called <u>diagonalizable</u>, if it admits reduction to diagonal form by similarity transformation to a diagonal matrix: $\exists X : A = XDX^{-1}$,

where X is nonsingular,
$D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_i \in \sigma(A)$

Ⓣ $A \in \mathbb{C}^{n\times n}$ has a diagonal form:
$A = XDX^{-1}$, $D = \mathrm{diag}(\sigma(A)) \Longleftrightarrow$
A has n lin. ind. eigenvectors $\Longleftrightarrow$
A is semisimple

Δ $A = XDX^{-1} \,\big|\, \cdot X$ Represent D, X as $X = \begin{pmatrix} | & & | \\ x_1 & \ldots & x_n \\ | & & | \end{pmatrix}$, $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$

$$AX = XD$$
$$\Downarrow$$
$A x_i = d_i x_i \Rightarrow$ x_i are eigenvectors of A
d_i are eigenvalues $\Rightarrow$
$\nu(\lambda) = \mu(\lambda)$, i.e. A is semisimple $\blacktriangle$

2) Jordan form

Reduction to a Jordan form is possible for $\forall A \in \mathbb{C}^{n\times n}$:
$\exists X : A = XJX^{-1}$, where J is block-diagonal matrix, consisting of Jordan boxes:

$$\underbrace{J = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_p \end{pmatrix}}_{n\times n} \qquad \overset{\text{Jordan box}}{\underset{\downarrow}{}} \quad J_i = \underbrace{\begin{pmatrix} J_{i1} & & \\ & \ddots & \\ & & J_{is_i} \end{pmatrix}}_{m_i \times m_i}$$

$$\underset{\substack{\uparrow \\ \text{Jordan cell}}}{J_{i_k}} = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{l_i \times l_i}$$

$m_i = \mu(\lambda_i)$ algebraic multiplicity

$s_i = \nu(\lambda_i)$ geometric multiplicity

$l_i \le m_i$, l_i is index of λ_i

Each Jordan cell J_{ik} corresponds to a different eigenvector associated with λ_i.

Ex:
$$\begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Jordan form is not used in numerical analysis, as it is sensitive to perturbations.

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

$$A = \begin{pmatrix} 2 & 1 \\ \varepsilon & 2 \end{pmatrix} \text{ where } \varepsilon = 1e-10$$
$$\Downarrow$$
$$J = \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix}$$

Matlab command: jordan(A)

Uniqueness of Jordan form

Since Jordan cells can be placed in any order, Jordan form is not unique.
Jordan decomposition is not unique:

$$A = X_1 J_1 X^{-1} = X_2 J_2 X_2^{-1}$$

⑥

Schur form

Reduction to a Schur form is possible for $\forall A \in \mathbb{C}^{n\times n}$: $\exists Q : A = QRQ^{-1}$, where Q is unitary $(Q^{-1} = Q^H)$, R is upper triangular, $\operatorname{diag}(R) = \sigma(A)$

Since Q is unitary, this is a unitary transformation $\boxed{A = Q R Q^H}$

R is called Schur form of A

$\operatorname{eig}(R) = \operatorname{eig}(A)$

Uniqueness of Schur decomposition

Since eigenvalues in R can be placed in different order, Schur decomposition is not unique

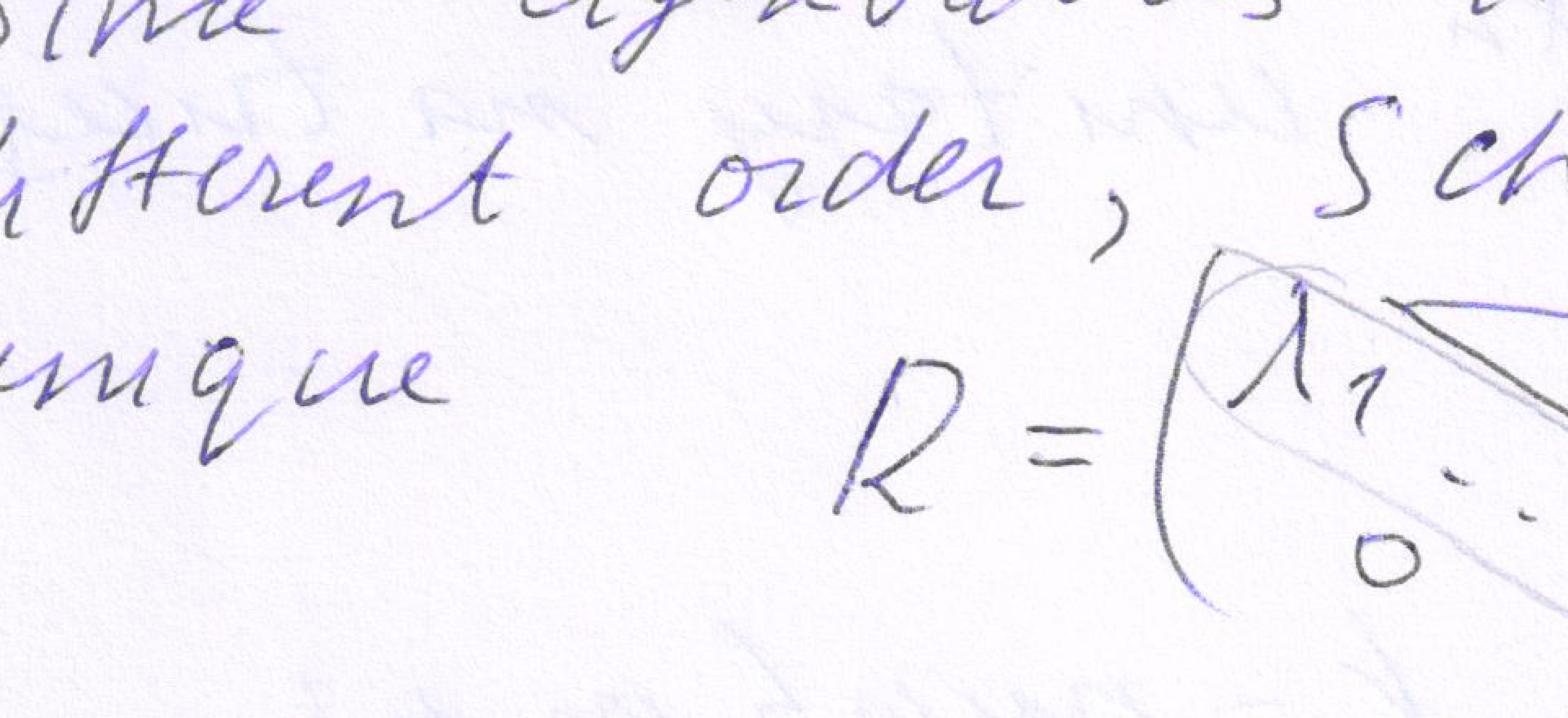
$$R = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

any order

Quasi-Schur (real Schur form)

When A is real, but $\sigma(A) \in \mathbb{C}$,
$A \in \mathbb{R}^{n\times n}$

there is a real Schur form:



$\times$ is 1×1 block for real λ_i

$\square$ is 2×2 block for complex conjugate pair

$\square = \begin{pmatrix} x & z \\ z & x \end{pmatrix}$, $x = \operatorname{Re}(\lambda)$ $\lambda_1 = x + iy$
$\lambda_2 = x - iy$

Matlab commands:

- real Schur: schur(A)
- complex schur: schur(A, 'complex')

4) Singular value decomposition (SVD)

⊜ Singular value of A

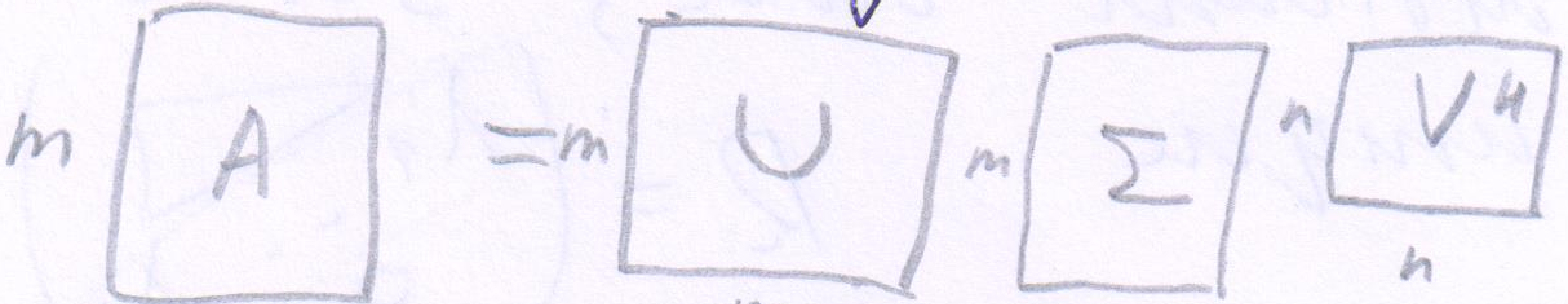
σ_i is called a singular value of A, if σ_i^2 is eigenvalue of AA^H

$\sigma_i = \sqrt{\lambda_i}\ ,\ \lambda_i \in \sigma(AA^H)$

Singular value decomposition is possible for $\forall\, A \in \mathbb{C}^{m \times n}$:

$\exists\, U \in \mathbb{C}^{m \times m},\ V \in \mathbb{C}^{n \times n}$ unitary matrices

$$\boxed{A = U \Sigma V^H}$$



$\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_k)\ ,\ k = \min\{m, n\}$

~~σ_i can be chosen all positive~~

σ_i can be placed in descending order:

$\sigma_1 \geq \sigma_2 \geq \dots \sigma_k \geq 0$

$$m > n: \quad \underbrace{\begin{pmatrix} \sigma_1 & \dots & 0 \\ 0 & & \sigma_n \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix}}_{n} \begin{matrix} \}\, n \\ \\ \}\, m \end{matrix}$$

$$m < n: \quad \underbrace{\begin{pmatrix} \sigma_1 & & 0 & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & \sigma_m & 0 & \dots & 0 \end{pmatrix}}_{n} \Bigg\}\, m \qquad \text{(first block: } \underbrace{\quad}_{m}\text{)}$$

8 Ⓣ $\|A\|_2 = \|\Sigma\|_2$

Relation between Schur and SVD

$$AA^H = U\Sigma V^H (U\Sigma V^H)^H = U\Sigma \underbrace{V^H V}_{I} \Sigma^H U^H =$$

$$= U \underbrace{\Sigma \Sigma^H}_{D} U^H = UDU^H \Rightarrow$$

D - diagonal matrix

D is Schur form of AA^H

Ⓣ If A is Hermitian ($A = A^H$), then its Schur form is diagonal

Δ If $A = A^H$, then AA^H is Hermitian:

$$(AA^H)^H = (A^H)^H A^H = AA^H$$

See PA 2-6 for illustration in Matlab: for a Hermitian positive definite matrix Schur form will be diagonal

PA 2-6. Compute Schur form for a Hermitian pos. def. A using svd function and not using schur function

---

PA 2-6: A is Hermitian pos. def. $\Rightarrow$

$A = LL^H$ (Cholessky decomposition)

Let $L = U\Sigma V^H \Rightarrow A = (U\Sigma V^H)(U\Sigma V^H)^H =$

$$= U\Sigma \underbrace{V^H V}_{I} \Sigma^H U^H = UDU^H \Rightarrow$$

$$A = UDU^H$$

SVD is connected with rank, kernel and range:

$A \in \mathbb{C}^{m\times n}$, $m \geq n$

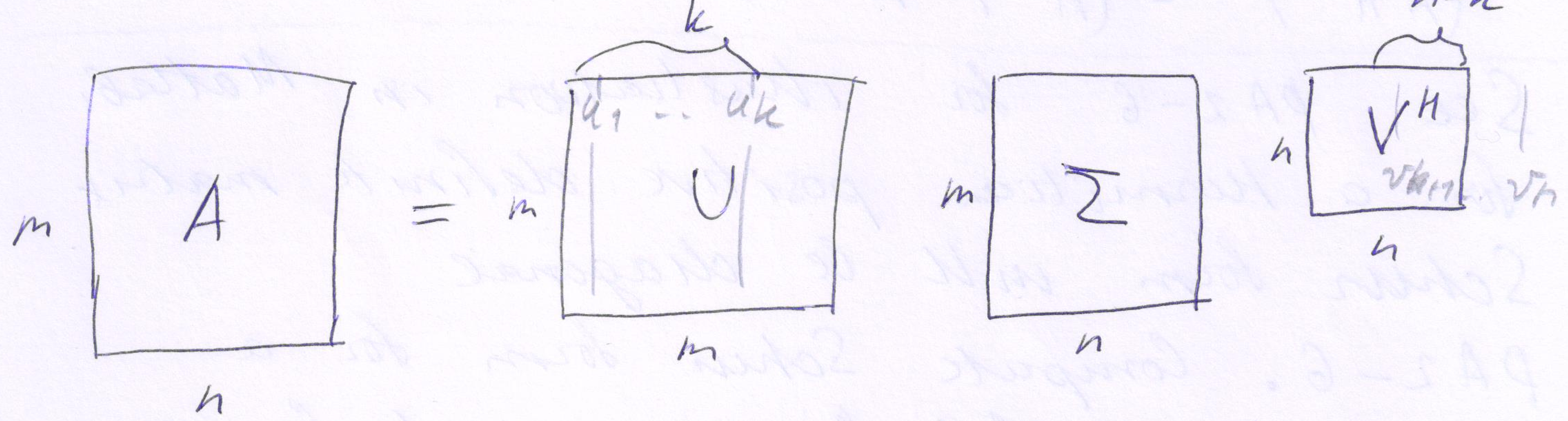
$$A = U\Sigma V^H, \quad \Sigma = \underbrace{\left(\begin{array}{ccccc} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ 0 & & & & \ddots \; 0 \\ \hline 0 & \cdots & & & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & & & 0 \end{array}\right)}_{n} \begin{array}{l} \left.\vphantom{\begin{array}{c} a\\a\\a\\a\\a \end{array}}\right\} n \\ \vphantom{\begin{array}{c} a\\a\\a \end{array}} \end{array} \Bigg\} m$$

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > \sigma_{k+1} = \dots = \sigma_n = 0$

$\text{rank}(A) = k$

$\text{Ran}(A) = \text{span}\{u_1, \dots, u_k\}$

$\text{Null}(A) = \text{span}\{v_{k+1}, \dots, v_n\}$



$\forall x \in \mathbb{C}^n$ $x = Vy$ - basis decomposition

V is a matrix of basis vectors, V is unitary,

y is a vector of coefficients.

$A = U\Sigma V^H \,|\, V$

$AV = U\Sigma \quad \Rightarrow \quad Ax = AVy = U\Sigma y \in$

$\text{span}\{\sigma_1 u_1, \dots, \sigma_n u_n\} =$

$= \text{span}\{\sigma_1 u_1, \dots, \sigma_k u_k\}$

⑩ LU-factorization

$A = LU$, L is lower triangular $L = (\Delta^{0})$

U is upper triangular $U = ({}_{0}\nabla)$

LU-factorization is not always possible

Ⓣ PLU-factorization

|| $\forall A$, A is nonsingular, $\exists P: A = P^{-1}LU$,
|| where P is permutation matrix

Permutation matrix is identity matrix with permuted columns

$$I = \begin{array}{ccc} \left(\begin{matrix} 1 & 0 & 0 \\ \ddots & 1 & \ddots \\ 0 & 0 & 1 \end{matrix}\right) \\ \begin{matrix} | & | & | \\ e_1 & e_i & e_n \end{matrix} \end{array}, \quad P \text{ is permutation of } \{e_i\}$$

e_1 e_i e_n unit orts

$P| \; A = P^{-1}LU$

$PA = LU$

PA are permuted rows of A

AP are permuted columns of A

Note that P is orthogonal: $P^{-1} = P^T$

PLU-factorization is called <u>factorization with partial pivoting</u>

Uniqueness of LU-factorization

Even when LU-factorization $\exists$, it is not unique: take $\{c_i\}$, $c_i \neq 0$ any set of numbers

Either $L = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$ or $U = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$,

numbers $\{c_i\}$ can be placed in any order.

Ⓣ || If $A = LU$ and $L = \begin{pmatrix} 1 & & 0 \\ \triangleright & \ddots & \\ & & 1 \end{pmatrix} \Rightarrow$
|| this factorization is unique

Matlab command: lu(A)

$[L, U] = lu(A)$

6) Cholesky factorization

When $A = A^H$, we can try to compute U as L^H: $A = \underset{\substack{\| \\ L}}{L}\,\underset{\substack{\| \\ U}}{L^H}$ to preserve symmetry

or $A = \underset{\substack{\| \\ L}}{R^H}\,\underset{\substack{\| \\ U}}{R}$

Cholesky factorization ~~of~~ for Hermitian A is not always possible

(T1) || If $\exists L: A = LL^H$, then L is unique.

(T2) || $\forall A \in \mathbb{C}^{n \times n}$ $\exists L: A = LL^H \Leftrightarrow A$ is
|| Hermitian positive definite

(T3) || Cholesky factorization is unique, if
|| all diagonal entries of R are positive

See PA3-1 for illustration.

PA3-1: For a Hermitian pos. def. A compute LU-factorization of A, where $l_{11} = l_{22} = \dots = l_{nn} = 1$?

12) A is Hermitian pos. def. =>

$A = \tilde{L}\tilde{L}^H$

find LU: $L = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ \triangle & & 1 \end{pmatrix}$

$$A = \tilde{L}\tilde{L}^H = \underbrace{\tilde{L}\,D^{-1}}_{\overset{\shortparallel}{L}}\underbrace{D\,\tilde{L}^H}_{\overset{\shortparallel}{U}}, \quad D = \mathrm{diag}(\tilde{L})$$

$$D = \begin{pmatrix} d_{11} & & & 0 \\ & d_{22} & & \\ & & \ddots & \\ 0 & & & d_{nn} \end{pmatrix} \Rightarrow D^{-1} = \begin{pmatrix} \frac{1}{d_{11}} & & & 0 \\ & \frac{1}{d_{22}} & & \\ & & \ddots & \\ 0 & & & \frac{1}{d_{nn}} \end{pmatrix}$$

$$\tilde{L}\,D^{-1} = L \Rightarrow \tilde{L} = \boxed{L}\,D$$

find L

<u>Matlab command:</u>

mrdivide or /

```
% XA = B
  X = A/B
R = chol(A) % gives upper-triangular part
L = chol(A, 'lower') %gives lower-triangular
                              part
```

Positive definite matrices

⊙ Real positive definite

$A \in \mathbb{R}^{n \times n}$

$(Au, u) > 0 \ \forall u \in \mathbb{R}^n, u \neq 0$

semidefinite: $(Au, u) \geq 0$

$\boxed{\text{Matlab: } u' * A * u}$

% $(Au, u) = u^H A u$

⊙ Symmetric positive definite

$A \in \mathbb{R}^{n \times n}, \ A = A^T$

$(Au, u) > 0 \ \forall u \in \mathbb{R}^n, u \neq 0$

⊙ Hermitian positive definite

$A \in \mathbb{C}^{n \times n}, \ A = A^H$

$(Au, u) > 0 \ \forall u \in \mathbb{C}^n, u \neq 0$

Inner product (Au, u) can be real only when $A = A^H$

Example Real pos. def. matrix can be nonsymmetric. Take block-diagonal A:

$$A = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

A is pos. def but $A \neq A^T$

Eigenvalues: $(2 \ 5 \ 2 \ 2)$ are all positive

Matlab: check positive definiteness by Cholessky decomposition

`chol(A)`

T1 Complex A is pos. def. $\Longleftrightarrow$
$A = A^H$ and $\sigma(A) > 0$ (eigenvalues are positive)

T2 $\forall A \in \mathbb{C}^{n\times n}$ there is a decomposition of A:
$A = H + iS$, where
$H = \frac{1}{2}(A + A^H)$, $S = \frac{1}{2i}(A - A^H)$
H, S are Hermitian
iS is skew-Hermitian
For eigenvalues of A holds:
1) $\lambda_{min}(H) \leq Re(\lambda_i) \leq \lambda_{max}(H)$
2) $\lambda_{min}(S) \leq Im(\lambda_i) \leq \lambda_{max}(S)$

T3 If A is real pos. def., then
1) $\exists A^{-1}$
2) $\exists \alpha > 0 : (Au, u) \geq \alpha \|u\|_2^2 \quad \forall u \in \mathbb{R}^n$

T4 $\forall A \in \mathbb{C}^{m\times n}$ $A^H A$ is Hermitian pos. semidefinite:
$\Delta \ (A^H A x, x) = (Ax, Ax) \geq 0 \quad \forall x \in \mathbb{C}^n$