Lecture 1 Fundamentals of Linear Algebra

Review of basic matrix theory, introduction of elementary notation

We consider $\mathbb{C}$-spaces.

- A complex matrix $A \in \mathbb{C}^{n \times m}$ is an array $n \times m$ of ~~com~~ $\mathbb{C}$-numbers

  a_{ij}, $j = \overline{1,m}$, $i = \overline{1,n}$ $\quad$ i – row, j = column

Main operations with matrices

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

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

- Multiplication by a scalar

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

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

- Multiplication by a matrix

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

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

- column-vector

  $$a_{*j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \quad j\text{-th column of } A$$

- row-vector

  $$a_{i*} = (a_{i1}, a_{i2}, \dots, a_{im}) \quad i\text{-th row of } A$$

2)

Transpose of a matrix

For $A \in \mathbb{C}^{n\times m}$ it's a matrix $C \in \mathbb{C}^{m\times n}$

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

denoted by A^T

Properties of matrix transposition

1) $(AB)^T = B^T A^T$
2) $(\lambda A)^T = \lambda A^T$
3) $(A^T)^T = A$
4) $(A+B)^T = A^T + B^T$

Transpose (complex) conjugate [git]

of a matrix A is

$$A^H \stackrel{def}{=} \bar{A}^T = \overline{A^T}$$

where the bar denotes element-wise complex conjugation. $z = x+iy$

$\bar{z} = x - iy$

Square matrices and eigenvalues

A matrix is called square if it has the same number of rows and columns:

$m=n \Rightarrow A \in \mathbb{C}^{n\times n}$

Identity matrix

$$I = \{\delta_{ij}\}_{i,j=\overline{1,n}} \qquad \delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$$ Kronecker delta.

Property of identity matrix:

$AI = IA \quad \forall A \in \mathbb{C}^{n\times n}$

Inverse matrix (of I)

C is called inverse to A, if $CA = AC = I$. Denoted by A^{-1}

Properties of matrix inverse:

1) $(AB)^{-1} = B^{-1}A^{-1}$
2) $(A^T)^{-1} = (A^{-1})^T$
3) $(kA)^{-1} = k^{-1}A^{-1}$

Determinant of a matrix

(recursive definition)

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j}), \text{ where}$$

$A \in \mathbb{C}^{n\times n}$, $A_{1j} \in \mathbb{C}^{(n-1)\times(n-1)}$ – matrix obtained by deletion the 1st row and j-th column of A.

~~Singular matrix~~

~~Matrix is called singular, if $\det(A) = 0$~~

Determinants of transpose and inverse matrices

$$\det A^{-1} = \frac{1}{\det A}$$

$$\det A^T = \det A$$

Properties of determinant, $A \in \mathbb{C}^{n\times n}$

1) $\det(AB) = \det(A)\det(B)$
2) $\det(A^T) = \det A$
3) $\det(\alpha A) = \alpha^n \det A$
4) $\det(\overline{A}) = \overline{\det(A)}$
5) $\det(I) = 1$

4) ⊖ <u>Singular matix</u>

A matrix is called <u>singular</u>, if $\det(A)=0$

Otherwise ($\det(A) \neq 0$) it's called <u>nonsingular</u>

⊖ <u>Characteristic polynomial of a matrix</u>

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

⊖ <u>Eigenvalues and eigenvectors</u>

$\mathbb{C}$-scalar λ is called an eigenvalue of the square matrix A, if

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

u is called <u>eigenvector</u> of A associated with λ.

⊖ <u>Spectrum</u>

$\sigma(A)$ - set of all eigenvalues

<u>Proposition 1</u>

|| $\lambda \in \sigma(A) \iff \det(\underset{p_A(\lambda)}{\underset{\parallel}{A}} - \lambda I) = 0$

For $A \in \mathbb{C}^{n \times n}$ there exists max n different eigenvalues

P2 || $\lambda = 0,\ \lambda \in \sigma(A) \iff A$ - singular (admits zero as eigenvalues)

P3 || A is nonsingular ($\det A \neq 0$) $\iff \exists A^{-1}$ (admits an inverse)

⊖ <u>Spectral radius</u>

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

⊜ Trace of a matrix $A \in \mathbb{C}^{n \times n}$

$tr(A) = \sum_{i=1}^{n} a_{ii}$ – sum of diagonal entries

(P) $tr(A)$ is equal to the sum of eigenvalues counted with their multiplicities as roots of $p_A(\lambda)$

(P) $\lambda \in \sigma(A) \Leftrightarrow \bar{\lambda} \in \sigma(A^H)$

⊜ Left and right ~~eigenvalues~~ eigenvectors

$Au = \lambda u$, u – right ~~eigenvalue~~ vectors

$v^H A = \lambda v^H$, v – left eigenvector

$$\left.\begin{aligned} (Au)^H &= (\lambda u)^H \\ u^H A^H &= \bar{\lambda} u^H \\ A^H v &= \bar{\lambda} v \end{aligned}\right\} \text{Equivalently}$$

Types of square matrices

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

(1) Symmetric matrices

$\boxed{A^T = A}$

(2) Hermitian matrices

$\bar{A}^T = \boxed{A^H = A}$

Properties of Hermitian matrices

1) $\det A \in \mathbb{R}$
2) A – Herm. $\Rightarrow A^{-1}$ – Herm. (if $\exists$)
3) A, B – Herm $\Rightarrow AB$ – Herm. $\Leftrightarrow AB = BA$

4) A, B – Herm. $\Rightarrow A+B$ – Herm.

3 <u>Skew-symmetric</u>

$A^T = -A$

4 <u>Skew-Hermitian</u>

$A^H = -A$

5 <u>Normal</u> $\boxed{A^H A = A A^H}$

6 Nonnegative $a_{ij} \geq 0, \forall i, j = \overline{1,n}$

nonpositive $a_{ij} \leq 0$

positive $a_{ij} > 0$

negative $a_{ij} < 0$

7 <u>Unitary</u> $\boxed{Q^H Q = I}$

For nonsquare:
$Q^H Q = I$ (of different
$Q Q^H = I$ size)

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

8 Unitary matrix $Q \in \mathbb{R}^{n \times n}$ is called <u>orthogonal</u>
(i.e. $Q^H Q$ is diagonal)

$A A^T = A^T A = I$ $\quad Q Q^T = Q^T Q = I$

$\boxed{A^{-1} = A^T}$ $\quad Q^{-1} = Q^T$

<u>Structures of square matrices with respect to the location of zero entries</u>

1) <u>Diagonal</u> $a_{ij} = 0 \quad \forall j \neq i$

$A = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$

2) <u>upper-triangular</u> $a_{ij} = 0, \forall i > j$

<u>lower-triangular</u> $a_{ij} = 0 \quad \forall i < j$

3) <u>upper bidiagonal</u> $a_{ij} = 0 \quad \forall j \neq i$ or $j \neq i+1$

lower bidiagonal $a_{ij} = 0 \quad \forall j \neq i$ or $j \neq i-1$

4) <u>Tridiagonal</u> matrix

$a_{ij} = 0 \quad \forall (i,j): |j-i| > 1$

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

5) <u>Banded</u> matrix

$a_{ij} \neq 0 \iff i - m_l \leq j \leq i + m_u$

where $m_l, m_u \geq 0, \in \mathbb{Z}$

$m_l + m_u + 1$ is a bandwidth of A

m_l is lower bandwidth: $a_{ij} = 0 \;\forall i > j + m_l$

m_u is upper bandwidth $a_{ij} = 0 \;\forall j > i + m_u$

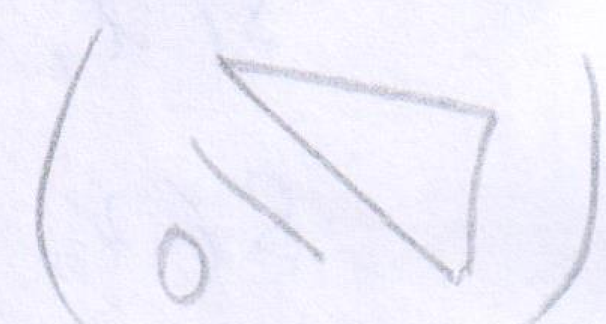
6) <u>Hessenberg</u> matrix (can be also nonsquare)

upper: $a_{ij} = 0 \quad \forall (i,j): i > j+1$

lower: $a_{ij} = 0 \;\forall (i,j): i < j+1$

$$\begin{matrix} X & X & 0 & 0 \\ X & X & X & 0 \\ X & X & X & X \\ X & X & X & X \end{matrix}$$

$$\begin{matrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{matrix}$$



7) <u>Permutation matrix</u>

obtained by permutation of columns of an identity matrix

8) <u>Block matrix</u>

$$A = \begin{pmatrix} A_{11} & \dots & A_{1l} \\ \vdots & & \\ A_{k1} & \dots & A_{kl} \end{pmatrix}$$

$k \times l$ – block sizes of A

9) <u>Block diagonal matrix</u>

$A = \text{diag}(A_{11}, A_{22}, \dots, A_{kk})$

10) <u>Block tridiagonal matrix</u>

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

8) <u>Inner product</u> скалярное произв-е

in $\mathbb{C}$-space $\mathcal{X}$ for $x \in \mathcal{X}$, $y \in \mathcal{X} \subset \mathbb{C}^n$

mapping $s: \mathcal{X} \times \mathcal{X} \to \mathbb{C}$

$s(x, y)$, which satisfies the following conditions:

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

$s(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 s(x_1, y) + \lambda_2 s(x_2, y)$

$\forall x_1, x_2 \in \mathcal{X}$, $\lambda_1, \lambda_2 \in \mathbb{C}$

2) $s(x,y)$ is Herm.

$s(y,x) = \overline{s(x,y)}$ $\forall x, y \in \mathcal{X}$

3) $s(x,y)$ is pos. definite

$s(x,x) > 0$ $\forall x \neq 0$

$s(x,x) \in \mathbb{R}$

$s(x,0) = s(x, 0 \cdot y) = 0 \cdot s(x,y) = 0$

$s(0,y) = 0$ $\forall y$

|| $s(x,x) \geq 0$ & $s(x,x) = 0 \iff x = 0$

т.к. $s(x,x) \in \mathbb{R}$

<u>Cauchy-Schwartz inequality:</u> нер-во Коши-Буняковского

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

Δ $s(x - \lambda y, x - \lambda y) = s(x,x) - \bar{\lambda} s(x,y) - \lambda s(y,x) + |\lambda|^2 s(y,y)$

for $y = 0$ inq. is satisfied trivially

for $y \neq 0$ we take $\lambda = \dfrac{s(x,y)}{s(y,y)} \Rightarrow$

$s(x - \lambda y, x - \lambda y) \geq 0$ by def.

$$S(x-\lambda y, x-\lambda y) = S(x,x) - 2\frac{|S(x,y)|^2}{S(y,y)} +$$

$$+ \frac{|S(x,y)|^2}{S(y,y)} = S(x,x) - \frac{|S(x,y)|^2}{S(y,y)} \geq 0 \Rightarrow$$

$$S(x,x)\,S(y,y) \geq |S(x,y)|^2$$ ▲

Eucledian inner product

If $X = \mathbb{C}^n$, then inner product $S(x,y)$ is called eucledian and denoted by (x,y).

$$(x,y) = \sum_{i=1}^{n} x_i \overline{y_i} \quad \text{or:} \quad (x,y) = y^H x$$

For $X = \mathbb{R}^n$ $\boxed{(x,y) = \sum_{i=1}^{n} x_i y_i}$ $\quad (x,x) = \sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + \ldots + x_n^2$

P $\boxed{(Ax,y) = (x, A^H y)}$ $\forall x,y \in \mathbb{C}^n,\ A \in \mathbb{C}^{n\times n}$

Property of unitary matrices: they preserve inner product:

$$(Qx, Qy) = (x,y) \quad \forall Q \text{ - unitary}$$

$$\Delta\ (Qx,Qy) = (x, \underbrace{Q^H Q}_{=I} y) = (x,y)$$ ▲

Vector norms

Norms are used to estimate errors in matrix computations.

Vector-norm on a $X \subset \mathbb{C}^n$ is a function $X \to \mathbb{R}$

denoted by $\|x\|$, $\|x\| \in \mathbb{R}$ $\forall x \in X$

which satisfies the following conditions:

10) 1) $\|x\| \geq 0 \quad \forall x \in X$

$\|x\| = 0 \iff x = 0$

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

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

triangle inequality

Particular cases of vector norms

- **Eucledian norm** $X = \mathbb{C}^n$

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

(P) $\| \; \|Qx\|_2 = \|x\|_2 \quad \forall x \in \mathbb{C}^n, \; Q$ - unitary

- **Hölder norms**

$$\|x\|_p \overset{def}{=} \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \quad \forall x \in \mathbb{C}^n$$

$$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n|$$

$$\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2}$$

$$\|x\|_\infty = \max_{i=\overline{1,n}} |x_i|$$

Cauchy - Schwartz inequality for Eucledian inner product and Eucledian norm

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

(P) $\|$ In finite space all norms are equivalent

Matrix norms are defined by vector norms.

Matrix norms

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

1) $\|A\| \geq 0 \quad \forall A \in \mathbb{C}^{n\times m}$
$\|A\| = 0 \iff A = 0$
2) $\|\alpha A\| = |\alpha| \cdot \|A\| \quad \forall A \in \mathbb{C}^{n\times m},\ \alpha \in \mathbb{C}$
3) $\|A+B\| \leq \|A\| + \|B\|$ triangle inequality

For matrices which admit matrix product there is property 4)

4) $\|AB\| \leq \|A\| \cdot \|B\|$ - multiplicative matrix norm

Often the term 'matrix norm' refers to operator matrix norm: the set of norms $\|\cdot\|_{pq}$ induced by two vector norms $\|\cdot\|_p$ and $\|\cdot\|_q$:

$$\|A\|_{pq} = \max_{\substack{x \in \mathbb{C}^m \\ x \neq 0}} \frac{\|Ax\|_p}{\|x\|_q}$$

When $p = q$

$$\|A\|_p = \max_{\|x\|_p = 1} \|Ax\|_p$$

Consistency of norms

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

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

Proposition

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

Proper matrix norms are operator norms.

12) Particular cases of matrix norms, which are not operator norms:

- <u>Hölder matrix norms</u>

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

- <u>Frobenius norm</u>: $p = 2$

$$\|A\|_F = \left(\sum_{j=1}^{m} \sum_{i=1}^{n} |a_{ij}|^2 \right)^{1/2}, \ A \in \mathbb{C}^{n \times m}$$

Frobenius norm is consistent with vector $\|\cdot\|_2$ norm $\quad \|A\|_F = \sqrt{tr(A^H A)} = \sqrt{tr(A A^H)}$

- <u>Euclidean norm</u> is Frobenius norm for square matrix $A \in \mathbb{C}^{n \times n}$

$$\|A\|_F = \|A\|_E = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

- <u>ℓ_1-norm</u>: $p = 1$

$$\|A\|_{\ell_1} = \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|, \ A \in \mathbb{C}^{n \times m}$$

- <u>M-norm</u>

$$\|A\|_M = n \max_{i,j=\overline{1,n}} |a_{ij}|, \quad A \in \mathbb{C}^{n \times n}$$

for square matrix

<u>Types of operator matrix norms</u>

1) <u>max column norm</u>

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

2) <u>max row norm</u>

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

3) Spectral norm

$\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A A^H)}$

It is known that $\sigma(A^H A) \geq 0$

$\|A\|_2 = \max\limits_{i=\overline{1,n}} \sigma_i$, where σ_i's is singular value

(!) σ_i is called a singular value of A, if σ_i^2 is an eigenvalue of $A^H A$

Propositions

P1. Matrix spectral norm $\|\cdot\|_2$ is consistent with vector norms $\|\cdot\|_2$, $\|\cdot\|_1$ with $\|\cdot\|_1$, $\|\cdot\|_\infty$ with $\|\cdot\|_\infty$.

P2. Matrix M-norm $\|\cdot\|_M$ is consistent with vector norms $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$

P3. Multiplicative property $\|AB\| \leq \|A\|\|B\|$ holds for any operator norm and for Frobenius norm.