Lecture 1

5.4 Inner product space:

Definition 5.4.1:
INNER PRODUCT SPACE: Let $\mathbf{x}, \mathbf{y} \in V$ then:

$$\begin{aligned} \langle \mathbf{x}, \mathbf{x} \rangle &\geq 0 \quad \text{iff}\, \mathbf{x} \neq \mathbf{0} \\ \langle \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{y}, \mathbf{x} \rangle \quad \forall \mathbf{x}, \mathbf{y} \in V \\ \langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle &= \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V \quad \alpha, \beta \in \mathbb{R} \end{aligned}$$

STANDARD INNER PRODUCT:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T, \mathbf{y}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i w_i \tag{5.4.1}$$

where $\mathbf{w}$ has positive entries, and w_i called WEIGHTS.

$$\begin{aligned} &\text{VECTOR SPACE}\, \mathbb{R}^{m\times n} \quad \text{For } A, B \in \mathbb{R}^{m\times n} \quad \langle A, B \rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ij} \\ &\text{VECTOR SPACE}\, C[a,b] \quad \text{For } f, g \in C[a,b] \quad \langle f, g \rangle = \int_a^b f(x)g(x)dx \\ &\text{VECTOR SPACE}\, P_n \quad x_1, \ldots, x_n \text{ distinct real numbers} \quad \langle p, q \rangle = \sum_{i=1}^{n} p(x_i)q(x_i) \end{aligned} \tag{5.4.2,3,5}$$

Norm, Phytagorean law:

NORM: length of a vector $\mathbf{v} \in V$ given by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

$\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ ORTHOGONAL if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$
FROBENIUS NORM: for $A \in \mathbb{R}^{m\times n}$ $\|A\|_f = \sqrt{(\langle A, A \rangle)}$

Theorem 5.4.1. PYTHAGOREAN LAW
$\mathbf{u}, \mathbf{v}$ orthogonal then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Projection:

Definition 5.4.2:
$\mathbf{u}, \mathbf{v} \in V$ and $\mathbf{v} \neq \mathbf{0}$ then:

$$\begin{aligned} &\text{SCALER PROJECTION OF } \mathbf{u} \text{ onto } \mathbf{v} \quad \alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|} \\ &\text{VECTOR PROJECTION OF } \mathbf{u} \text{ onto } \mathbf{v} \quad \mathbf{p} = \alpha \left(\frac{1}{\|\mathbf{v}\|} \mathbf{v} \right) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \end{aligned} \tag{5.4.7}$$

Observations:
$\mathbf{v} \neq \mathbf{0}$ and $\mathbf{p}$ projection $\mathbf{u}$ onto $\mathbf{v}$ then:
1) $\mathbf{u} - \mathbf{p}$ and $\mathbf{p}$ orthogonal. 2) $\mathbf{u} = \mathbf{p}$ iff $\mathbf{u}$ scaler multiple $\mathbf{v}$
PROOFS: in book.

Cauchy-Schwarz inequality:

Theorem 5.2: CAUCHY SCHWARTZ INEQUALITY:
$\mathbf{u}, \mathbf{v} \in V$ then

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\| \tag{5.4.8}$$

iff $\mathbf{u}, \mathbf{v}$ linearly dependent.
PROOF:

$$1)\ \mathbf{v} = \mathbf{0} \Rightarrow |\langle \mathbf{u}, \mathbf{v} \rangle| = 0 = \|\mathbf{u}\| \|\mathbf{v}\|$$
$$2)\ \& v \neq \mathbf{0} \Rightarrow \mathbf{p} \perp \mathbf{u} - \mathbf{p} \Rightarrow \|\mathbf{p}\|^2 + \|\mathbf{u} - \mathbf{p}\|^2 = \|\mathbf{u}\|^2 \Rightarrow \|\mathbf{p}\|^2 = \|\mathbf{u}\|^2 - \|\mathbf{u} - \mathbf{p}\|^2$$
$$(\langle \mathbf{u}, \mathbf{v} \rangle^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{p}\|^2 \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \tag{5.4.9}$$

We see that (5.4.9) holds iff $\mathbf{u} = \mathbf{p}$ so if $\mathbf{v} = \mathbf{0}$ or $\mathbf{u}$ multiple of $\mathbf{v}$, so when $\mathbf{u}$ and $\mathbf{v}$ linearly independent.
consequence:
$\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$ so then

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1 \Rightarrow \exists \text{unique} \theta \in [0, \pi] \text{s.t.} \quad \cos(\theta) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \tag{5.4.10}$$

Normed linear space:

Definition 5.4.3:
V NORMED LINEAR SPACE if $\mathbf{v} \in V$ there is a associatedNORM $\|\mathbf{v}\| \in \mathbb{R}$ satisfying

$$\begin{aligned} &1) \|\mathbf{v}\| \geq 0 \quad \text{equility iff } \mathbf{v} = \mathbf{0} \\ &2) \|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\| \quad \text{for any } \alpha \in \mathbb{R} \\ &3) \|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \quad \forall \mathbf{v}, \mathbf{w} \in V \end{aligned}$$

Number 3 is called TRIANGLE INEQUALITY

Theorem 5.4.3
V inner product space then the following equation defines a norm on V

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \quad \text{for all } \mathbf{v} \in V$$

Definition 4:
$\mathbf{x}, \mathbf{y}$ vectors. Then distance between $\mathbf{x}$and $\mathbf{y}$ defined by $\|\mathbf{y} - \mathbf{x}\|$

Lecture 2

Orthogonal and orthonormal:

Parallelogram on V iff: $2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2, \forall \mathbf{u}\&\mathbf{v} \in V$
$\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$:
Orthogonal set: iff $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$,for all $i \neq j$
Definition 5.5.1
textscOrthonormal set: orthogonal set with extra condition: $\|\mathbf{v}_i\| = 1$ for all i
Notation: Kronecker delta symbol: $\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$
Orthonormal set of vectors: orthogonal set of unit vectors.

theorem 5.5.1:
$\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ orthogonal set, then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ linearly independent.
Proof:
$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = 0$ implies $c_1, \ldots, c_n = 0$ Take $j \in \{1, \ldots, n\}$ then we have:
$0 = \langle c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n, v_j \rangle = c_1\langle \mathbf{v}_1, \mathbf{v}_j \rangle + \ldots + c_n\langle \mathbf{v}_n, \mathbf{v}_j \rangle$
All zero ,because orthogonal except $\langle \mathbf{v}_j \mathbf{v}_j \rangle$
So we know that $0 = c_j\langle \mathbf{v}_j, \mathbf{v}_j \rangle$ so $c_j = 0$ so lin. independent.

B Orthonormal basis S if $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ and $S = \text{Span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$

Theorem 5.5.2:

$$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\} \text{ orhtonormal basis } V. \text{ If } \mathbf{v} = \sum_{i=1}^{n} c_i\mathbf{u}_i \Rightarrow c_i = \langle \mathbf{v}, \mathbf{u}_i \rangle$$

$$\text{Proof}\langle \mathbf{v}, \mathbf{u}_i \rangle = \left\langle \sum_{j=1}^{n} c_j\mathbf{u}_j, \mathbf{u}_j \right\rangle = \sum_{j=1}^{n} c_j\langle \mathbf{u}_j, \mathbf{u}_i \rangle = \sum_{j=1}^{n} c_j\delta_{ji} = c_i$$

Corollary 5.5.3:
Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ orthonormal basis V then:

$$\mathbf{u} = \sum_{i=1}^{n} a_i\mathbf{u}_i \quad \& \quad \mathbf{v} = \sum_{i=1}^{n} b_i\mathbf{u}_i \to \langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} a_ib_i$$

Proof:
By theorem 5.5.2 $\langle \mathbf{v}, \mathbf{u}_i \rangle = b_i \quad i = 1, \ldots, n$ so therefore

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \sum_{i=1}^{n} a_i\mathbf{u}_i, \mathbf{v} \right\rangle = \sum_{i=1}^{n} a_i\langle \mathbf{u}_i, \mathbf{v} \rangle = \sum_{i=1}^{n} a_i\langle \mathbf{v}, \mathbf{u}_i \rangle = \sum_{i=1}^{n} a_ib_i$$

Corollary 5.5.4: Formula of parseval

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_n\} \text{ orthonormal basis } V \,\&\, \mathbf{v} = \sum_{i=1}^{n} c_i\mathbf{u}_i \in V \Rightarrow \|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle = \sum_{i=1}^{n} c_i^2$$

Definition 5.5.2: $Q \in \mathbb{R}^{n\times n}$ Orthogonal matrix if column vectors Q form orthonormal set in $\mathbb{R}^n$

Theorem 5.5.5:
$Q \in \mathbb{R}^{n\times n}$ orthogonal iff $Q^TQ = I$
PROOF:
Column vectors: $\mathbf{q}_i^T\mathbf{q}_j = \delta_{ij}$. Because $\mathbf{q}_i^T\mathbf{q}_j$ is (i,j) entry of $Q^TQ \Rightarrow Q$ orthogonal $\Leftrightarrow Q^TQ = I$
Properties orthogonal matrices:
$Q \in \mathbb{R}^{n\times n}$ then

$a)$ column vectors Q orthonormal basis $\mathbb{R}^n$
$b)$ $Q^TQ = I$
$c)$ $Q^T = Q^{-1}$
$d)$ $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$
$e)$ $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$

PERMUATION MATRIX: P by reodering the columns of I in the order $(k_1,\dots,k_n)$ then $P = (\mathbf{e}_{k_1},\dots,\mathbf{e}_{k_n})$.
If $A \in \mathbb{R}^{m\times n}$ then

$$AP = (A\mathbf{e}_{k_1},\dots,A\mathbf{e}_{k_n}) = (\mathbf{a}_{k_1},\dots,\mathbf{a}_{k_n})$$

LEAST SQUARE SOLUTION:
$A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.

Definition 1 find $\hat{\mathbf{x}} \in \mathbb{R}^n$ s.t. $\|A\hat{\mathbf{x}} - \mathbf{b}\|^2$ is minimal
Definition 2 minimize $\mathbf{r} : \mathbb{R}^n \mapsto [0,\infty]$ where $\mathbf{r} := \|A\mathbf{x} - \mathbf{b}\|^2$
Definition 3 find orthogonal projection $\mathbf{p}$ of $\mathbf{b}$ onto subspace $R(A)$

NORMAL EQUATION: $A^T\mathbf{b} - A^TA\hat{\mathbf{x}} = 0$
Derive normal equations: $\forall x$

$$\mathbf{b} - A\hat{\mathbf{x}} \perp R(A) \Leftrightarrow (A\mathbf{x})^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \Leftrightarrow \mathbf{x}^T(A^T\mathbf{b} - A^TA\hat{\mathbf{x}}) = 0 \Leftrightarrow A^T\mathbf{b} - A^TA\hat{\mathbf{x}} = 0$$

Theorem 5.5.6:
If column vectors A orthonormal set then $A^TA = I$ and solution least squares problem is

$$\hat{\mathbf{x}} = A^t\mathbf{b}$$

PROOF:
Column vectors A orthonormal $\Rightarrow A^TA = I$ so then

$$A^T\mathbf{b} - A^TA\hat{\mathbf{x}} = 0 \Leftrightarrow A^T\mathbf{b} - I\cdot\hat{\mathbf{x}} = 0 \Leftrightarrow \hat{\mathbf{x}} = A^T\mathbf{b}$$

Theorem 5.5.7:
S subspace V and $\mathbf{x} \in V$. Let $\{\mathbf{u}_1,\dots,\mathbf{u}_n\}$ orthonormal basis S.If

$$\mathbf{p} = \sum_{i=1}^{n} c_i\mathbf{u}_i \tag{5.5.3}$$

where

$$c_i = \langle \mathbf{x}, \mathbf{u}_i\rangle \quad \forall i \tag{5.5.4}$$

then $\mathbf{p} - \mathbf{x} \in S^\perp$
PROOF:
By using the definitions (see book)

Lecture 3

Theorem 5.5.8:
Let $\mathbf{p}$ be the element in S closest to $\mathbf{x}$. So

$$\|\mathbf{y}-\mathbf{x}\| > \|\mathbf{p}-\mathbf{x}\| \quad \text{for any } \mathbf{y} \neq \mathbf{p} \in S$$

Proof:
$\mathbf{y} \in S$ and $\mathbf{y} \neq \mathbf{p}$ then

$$\begin{aligned} &\|\mathbf{y}-\mathbf{x}\|^2 = \|(\mathbf{y}-\mathbf{p}) + (\mathbf{p}-\mathbf{x})\|^2 \\ &\text{since } \mathbf{y}-\mathbf{p} \in S \text{ and Theorem 5.5.7+ Pythagorean Law} \\ &\|\mathbf{y}-\mathbf{x}\|^2 = \|\mathbf{y}-\mathbf{p}\| + \|\mathbf{p}-\mathbf{x}\|^2 > \|\mathbf{p}-\mathbf{x}\|^2 \end{aligned}$$

Corollary 5.5.9:
S nonzero subspace $\mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^m$. If $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ orthonormal basis S and $U = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$ then projection of $\mathbf{b}$ onto S given by

$$\mathbf{p} = UU^T\mathbf{b}$$

Proof:
By theorem 5.5.7:

$$\mathbf{p} = c_1\mathbf{u}_1 + \ldots + c_k\mathbf{u}_k = U\mathbf{c} \text{ where } \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{b} \\ \vdots \\ \mathbf{u}_k^T\mathbf{b} \end{bmatrix} = U^T\mathbf{b}$$

Gram-Schmidt Proces:

Theorem 5.6.1:

$$\begin{aligned} &\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \text{ basis } V \\ &\text{define } \mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1 \\ &\text{define recursively } \mathbf{u}_2, \ldots, \mathbf{u}_n : \mathbf{u}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1}-\mathbf{p}_k\|}(\mathbf{x}_{k+1}-\mathbf{p}_k) \quad \text{for } k = 1, \ldots, n-1 \\ &\text{where} \mathbf{p}_k = \langle \mathbf{x}_{k+1}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \ldots + \langle \mathbf{x}_{k+1}, \mathbf{u}_k \rangle \mathbf{u}_k \end{aligned}$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ orthonormal basis V
Proof:
By induction.

Lecture 4

Applications of GS-proces:

Theorem 5.6.2:
$A \in \mathbb{R}^{m\times n}$ of rank n then A can be factored in $A = QR$ where $Q \in \mathbb{R}^{m\times n}$ orthogonal and $R \in \mathbb{R}^{n\times n}$ uper triangle with positive diagonal entries.
Note: R nonsingular since $\det(R) > 0$
PROOF:
By Gram-Schmidt Proces and induction.

Theorem 5.6.3:
If $A \in \mathbb{R}^{m\times n}$ and of rank n then least square solution by $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$.
We may obtain $\hat{\mathbf{x}}$ by backsubstubtitution to solve $R\mathbf{x} = Q^T\mathbf{b}$
PROOF:

$$A\mathbf{x} = \mathbf{b} \Leftrightarrow A^TA\mathbf{x} = A^T\mathbf{b} \Leftrightarrow (QR)^TQR\mathbf{x} = (QR)^T\mathbf{b}$$
$$R^T(Q^TQ)R\mathbf{x} = R^TQ^T\mathbf{b} \Leftrightarrow R^TR\mathbf{x} = R^TQ^T\&b \Leftrightarrow R\mathbf{x} = Q^T\mathbf{b} \Leftrightarrow \mathbf{x} = R^{-1}Q^T\mathbf{b} = \hat{\mathbf{x}}$$

Lecture 5

Definition 6.1.1:
$A \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$ is EIGENVALUE/CHARACTERISTIC VALUE if exists $\mathbf{x} \in \mathbb{R}^n$ s.t. $A\mathbf{x} = \lambda\mathbf{x}$. Then $\mathbf{x}$ is called EIGENVECTOR/CHARACTERISTIC VECTOR belonging to λ.
Equivalent:
$A \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$

1) λ eigenvalue A
2) $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has nontrivial solution
3) $N(A - \lambda I) \neq \{\mathbf{0}\}$
4) $A - \lambda I$ singular
5) $\det(A - \lambda I) = 0$

Some facts:

1) $\lambda \in \mathbb{C}$eigenvalue $\Leftrightarrow \overline{\lambda} \in \mathbb{C}$eigenvalue
REASON: $A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} \Leftrightarrow A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$
2) $\det(A) = p(0) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n$
REASON see equation 6.1.4 till 6.1.6
3) $\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$
REASON Point 2

$B \in \mathbb{R}^{n\times n}$ is SIMILAR to $A \in \mathbb{R}^{n\times n}$ if there exists nonsingular S s.t. $B = S^{-1}AS$
Theorem 6.1.1:
$A, B \in \mathbb{R}^{n\times n}$. B similar to A if they have the same characteristic polynomial and therefore the same eigenvalues.
PROOF:

$$p_B(\lambda) = \det(B - \lambda I) = \det\big(S^{-1}AS - \lambda I\big) = \det\big(S^{-1}(A - \lambda I)S\big) = \det\big(S^{-1}\big)\det(A - \lambda I)\det(S) = p_A(\lambda)$$

Theorem 6.3.1:
$\lambda_1, \ldots, \lambda_k$ distinct eigenvalues of $A \in \mathbb{R}^{n\times n}$ then corresponding eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent.
PROOF: See book.
Definition 6.3.1:
$A \in \mathbb{R}^{n\times n}$ DIAGONALIZABLE if $\exists$nonsingular $X \in \mathbb{R}^{n\times n}$&diagonal $D \in \mathbb{R}^{n\times n}$ s.t. ,$x^{-1}AX = D$
Then X diagonalizes A
Theorem 6.3.2:
$A \in \mathbb{R}^{n\times n}$ diagonalizable $\Leftrightarrow$ A has n lin. independent eigenvectors.
PROOF: See book, there are also remarks there.

Lecture 6

Complex plane:

FIELD: a set K satisfies axioms.
Today $K = \mathbb{C}$ so:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \quad \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \quad \text{with } z_i, w_i \in \mathbb{C} \text{ so then } \sum_{i=1}^{n} \overline{z}_i w_i = \overline{\mathbf{z}}^T \mathbf{w} \in \mathbb{C}$$

$$\text{when } \mathbf{z} = a + bi \Rightarrow \overline{\mathbf{z}} = a - bi$$

$$\overline{\mathbf{z}}^T = \mathbf{z}^* = \mathbf{z}^H \quad \text{called Hermitian transpose}$$

$\|\mathbf{z}\| = \sqrt{\mathbf{z}^H \mathbf{z}}$
Definition 6.4.1:
V inner product space $\mathbf{w}, \mathbf{z} \in V$ then $\langle \mathbf{z}, \mathbf{w} \rangle$:

$$1)\langle \mathbf{z}, \mathbf{z} \rangle \geq 0 \quad \text{equility iff } \mathbf{z} = \mathbf{0}$$

$$2)\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle} \quad \forall \mathbf{z}, \mathbf{w} \in V$$

$$3)\langle \alpha\mathbf{z} + \beta\mathbf{w}, \mathbf{u} \rangle = \alpha\langle \mathbf{z}, \mathbf{u} \rangle + \beta\langle \mathbf{w}, \mathbf{u} \rangle$$

We also see that $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z}$

Hermitian matrix:

$M = (m_{ij}) \in \mathbb{C}^{m \times n}$ with $m_{ij} = a_{ij} + b_{ij}$ so then we have that $M = A + iB \Rightarrow \overline{M} = A - iB$.
For matrices we have the following rules:

$$1)(A^H)^H = A \quad 2)(\alpha A + \beta B)^H = \overline{\alpha} A^H + \overline{\beta} B^H \quad 3)(AC)^H = C^H A^H$$

Definition 6.4.2:
Matrix M is HERMITIAN if $M = M^H$

Theorem 6.4.1:
The eigenvalues of Hermitian matrix are real. Furthermore, eigenvectors belonging to distinct eigenvalues are orthogonal.
PROOF: See book.

Unitary matrices and diagonalizibility:

Definition 6.4.3:
$n \times n$ matrix U is UNITARY if column vectors form orthonormal set in $\mathbb{C}^n$.
Theorem (numberless):
U is unitary $\Leftrightarrow U^H U = I_{n \times n} \Leftrightarrow U$ is nonsingular and $U^{-1} = U^H$
Corollary 6.4.2:
Eigenvalues Hermitian A distinct $\Rightarrow \exists$ unitary U that diagonalizes A
PROOF: See book.

Theorem 6.4.3:SCUR'S THEOREM
For each $n \times n$ matrix A there exists unitary U s.t. $U^H AU$ is upper triangular.
PROOF: Really long, by induction, see book.

Lecture 7

Theorem 6.4.4:
If A Hermitian then there exists unitary U that diagonalizes A.
PROOF:
By theorem 6.4.3. exists U s.t. $U^H AU = T$. Furthermore:

$$T^H = (U^H AU)^H = U^H A^H U = U^H AU = T$$

So T is Hermitian and therefore diagonal.

Definition 6.4.4:
subspace $S \subset \mathbb{R}^n$ is INVARIANT under A if $\mathbf{x} \in S \Rightarrow A\mathbf{x} \in S$ for each $\mathbf{x}$

Lemma 6.4.5:
$A \in \mathbb{R}^{n\times n}$ with $\lambda_1 = a + bi$, with $a, b \in \mathbb{R}$ and $b \neq 0$.
Let $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ belonging to λ_1. If $S = \text{span}(\mathbf{x}, \mathbf{y})$ then $\dim S = 2$ and S is invariant under A.
PROOF:

$$\begin{aligned}
&A\mathbf{z}_1 = \lambda_1\mathbf{z}_1\\
&A\mathbf{z}_1 = A\mathbf{x} + iA\mathbf{y}\\
&\lambda_1\mathbf{z}_1 = (a + bi)(\mathbf{x} + i\mathbf{y}) = (a\mathbf{x} - b\mathbf{y}) + i(b\mathbf{x} + a\mathbf{y})\\
&\Rightarrow A\mathbf{x} = a\mathbf{x} - b\mathbf{y} \quad A\mathbf{y} = b\mathbf{x} + a\mathbf{y}\\
&\mathbf{w} = c_1\mathbf{x} + c_2\mathbf{y} \in S\\
&A\mathbf{w} = c_1(a\mathbf{x} - b\mathbf{y}) + c_2(b\mathbf{x} + a\mathbf{y}) = (c_1a + c_2b)\mathbf{x} + (c_2a - c_1b)\mathbf{y}\\
&\Rightarrow A\mathbf{w} \in S
\end{aligned}$$

Theorem 6.4.6: Real schur decomposition:
Let $A \in \mathbb{R}^{n\times n}$. Exists orthogonal $Q \in \mathbb{R}^{n\times n}$ and "quasi upper triangle $T \in \mathbb{R}^{n\times n}$ s.t $A = QTQ^T$

$$T = \begin{pmatrix} B_1 & \dots & \dots & \dots \\ 0 & B_2 & \dots & \dots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & B_k \end{pmatrix}$$

B_i's are 2×2 or 1×1 and determined as follows:
Compute all eigenvalues $\lambda_1, \dots, \lambda_n$ of A
Suppose $\lambda_1, \dots, \lambda_r$ not real and $\lambda_{r+1}, \dots, \lambda_n$ real.
$\lambda_1, \dots, \lambda_r$ appear in complex conjugate pairs, say:
$\lambda_1, \overline{\lambda_1}, \lambda_2, \overline{\lambda_2}, \dots, \lambda_{\frac{r}{2}}, \overline{\lambda_{\frac{r}{2}}}$
Suppose $\lambda_j = a_j + ib_j$ and $\overline{\lambda_j} = a_j - ib_j$ This gives $\frac{r}{2}$, 2×2 matrices $B_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$
The remaining real $\lambda_{r+1}, \dots, \lambda_n$ give $n - r$,1×1 matrices namely $B_j = \lambda_j$
PROOF: skipped.

Corollary 6.4.7: SPECTRAL THEOREM FOR REAL SYMMETRIC MATRICES
A real symmetric, then $\exists$ orthogonal Q diagonalizes A:
$Q^T AQ = D$ where D diagonal.

A is SKEW HERMITIAN: if $A^H = -A$

Definition 6.4.5:
A is NORMAL if $AA^H = A^H A$.
How to derive this?

$$
\begin{aligned}
&A^H = UD^H U^H \Rightarrow A = UDU^H \\
&AA^H = UDU^H UD^H U^H = UDD^H U^H \\
&A^H A = UD^H U^H UDU^H = UD^H DU^H \\
&\text{since } D^H D = DD^H = \begin{bmatrix} |\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \vdots & \\ & & & |\lambda_n|^2 \end{bmatrix} \Rightarrow AA^H = A^H A
\end{aligned}
$$

A matrix has a complete orthonormal set of eigenvectors iff it is normal.

Theorem 6.4.8:
A is normal iff A possesses a complete orthonormal set of eigenvectors:
PROOF: See book.

Lecture 8

Theorem 6.5.1 if A is $m \times n$ then A has SVD

Proof:

Long proof, intermediate facts and their proofs.

Step 1: Finding singular values

PROOF:

$\boxed{A}$ $\lambda_i \in \mathbb{R}$ for $i = 1, \ldots, n$.fact $\lambda_i \geq 0$

PROOF $A^T A\mathbf{x}_i = \lambda_i \mathbf{x}_i \Rightarrow \mathbf{x}_i^T A^T A\mathbf{x}_i = \lambda_i \|\mathbf{x}_i\|^2 \Rightarrow \|A\mathbf{x}_i\|^2 = \lambda_i \|\mathbf{x}_i\|^2$

$\Rightarrow \lambda_i \geq 0$

$\boxed{B}$ Relabel eigenvalues:$\lambda_1 \geq \ldots \geq \lambda_n$.fact rank$(A) = r \Rightarrow \lambda_{r+1} = \ldots = 0$

PROOF V nonsingular and orthogonal $\Rightarrow A^T A = VDV^T \Rightarrow VDV^T = D \Rightarrow \text{rank}\,(A^T A) = \text{rank}\,(D)$

also $\text{rank}(AA^T) = n - \dim \mathcal{N}(A^T A) = n - \dim \mathcal{N}(A) = \text{rank}(A)$

$\Rightarrow \text{rank}(A^T A) - \text{rank}(D) = \text{Rank}(A) = r$

$\Rightarrow \lambda_1 \geq \ldots \geq \lambda_r > 0 \quad \lambda_{r+1} = \ldots = \lambda_n = 0$

$\boxed{C}$ $\sigma_1 \geq \ldots \geq \sigma_r > 0 \& \sigma_{r+1} = \ldots = \sigma_n = 0$

Proof $\sigma_i = \sqrt{\lambda_i} \quad$ for $i = 1, \ldots, n$

$\sigma_1 \geq \ldots \geq \sigma_r > 0 \quad \sigma_{r+1} = \ldots = \sigma_n = 0$

Step 2: Finding orthogonal V

$$\text{Let } V_1 = (\mathbf{v}_1, \ldots, \mathbf{v}_r) \quad V_2 = (\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n) \quad \Sigma_1 = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix} \quad \Sigma = \left(\begin{smallmatrix} \Sigma_1 & O \\ O & O \end{smallmatrix}\right)$$

$\boxed{A}$ $AV_2 = 0$

PROOF $V\Sigma \Leftrightarrow A^T A \left(V_1 \; V_2 \right) = \left(V_1 \; V_2 \right) \left(\begin{smallmatrix} \Sigma_1 & O \\ O & O \end{smallmatrix}\right) \Leftrightarrow \left(A^T AV_1 \mid A^T AV_2 \right) = \left(V_1 \Sigma_1 \mid 0 \right)$

$A^T AV_2 = 0 \Rightarrow V_2^T AV_2 = 0$.Let $\mathbf{x} \in \mathbb{R}^{n-r}$

$\Rightarrow \mathbf{x}^T V_2^T A^T V_2 \mathbf{x} = 0 \Leftrightarrow (AV_2\mathbf{x})^T (AV_2\mathbf{x}) = 0 \Leftrightarrow \|AV_2\mathbf{x}\|^2 = 0$

$\mathbf{x}$ arbitrary $\Rightarrow AV_2 = 0$

$\boxed{B}$ $AV_1 V_1^T A = A$

PROOF $VV^T = I \Rightarrow \left(V_1 \; V_2 \right) \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = I \Rightarrow V_1 V_1^T + V_2 V_2^T = I$

$\Rightarrow AV_1 V_1^T = A(I - V_2 V_2^T) = A \Rightarrow AV_1 V_1^T = A$

Step 3: Finding an orthogonal U

$$\Sigma = \left(\begin{smallmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{smallmatrix}\right) \quad U = \left(U_1 \; U_2 \right) \quad \text{where}\, U_1 = m \times r \& U_2 = m \times (m-r)$$

$$U_1 = \left(\mathbf{u}_1 \cdots \mathbf{u}_m \right) \text{with}\, u_i = \frac{1}{\sigma_i} A\mathbf{v}_i$$

$\boxed{A}$ $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ orthonormal set in $\mathbb{R}^m$

Proof $\mathbf{u}_i^T \mathbf{u}_i = \left(\frac{1}{\sigma_i} A\mathbf{v}_i\right)^T \frac{1}{\sigma_i} A\mathbf{v}_i = \frac{1}{\sigma_i^2} \mathbf{v}_i^T A^T A\mathbf{v}_i.$

let $A^T A\mathbf{v}_i = \lambda \mathbf{v}_i = \sigma_i^2 \mathbf{v}_i \Rightarrow \mathbf{u}_i^T \mathbf{u}_i = 1$

$$i \neq j \Rightarrow \mathbf{u}_i^T \mathbf{u}_j = \left(\frac{1}{\sigma_i} A\mathbf{v}_i\right)^T \frac{1}{\sigma_j} A\mathbf{v}_j = \frac{1}{\sigma_i \sigma_j} \mathbf{v}_i^T A^t A\mathbf{v}_j = \frac{\sigma_j^2}{\sigma_i \sigma_j} \mathbf{v}_i^T \mathbf{v}_j = 0$$

Step final: Final conclusion:

We have:

$$U_i \Sigma = \left(\tfrac{1}{\sigma_1} A\mathbf{v}_1 \cdots \tfrac{1}{\sigma_r} A\mathbf{v}_r \right) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} = \left(A\mathbf{v}_1 \; A\mathbf{v}_2 \cdots A\mathbf{v}_r \right) = AV_1$$

Take U_2 s.t. $(U_1|U_2)$ orthogonal by Gramm schmidt.
When we calculate $U\Sigma V^T$ we find:

$$\left(U_1 \; U_2 \right) \left(\begin{smallmatrix} \Sigma_1 & O \\ O & O \end{smallmatrix}\right) \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 \Sigma_1 V_1^T = AV_1 V_1^T = A$$

Observations about SVD:

In this little part I will use ONB for Orthonormal Basis.

1) $\sigma_1, \ldots, \sigma_n$ of A unique U, V not
2) V diagonalizes $A^T A \Rightarrow \mathbf{v}_j$ eigenvectors $A^T A$
3) $AA^T = U\Sigma\Sigma^T U^T \Rightarrow U$ Diagonalizes AA^T and $\mathbf{u}_j$ eigenvectors AA^T

4a) $AV = U\Sigma \Rightarrow A\mathbf{v}_j = \sigma_j \mathbf{u}_j$ for $j = 1, \ldots, n$

4b) $A^T U = V\Sigma^T \Rightarrow A^T \mathbf{u}_j = \sigma_j \mathbf{v}_j$ for $j = 1, \ldots, n . A^T \mathbf{u}_j = \mathbf{0}$ for $j = n+1, \ldots, m$

5) $\text{rank}(A) = r \Rightarrow a)\, \mathbf{v}_1, \ldots, \mathbf{v}_r$ ONB $R(A^T)$ & $b)\mathbf{v}_{r+1}, \ldots, v_n$ ONB $\mathcal{N}(A)$
$c)\, \mathbf{u}_1, \ldots, \mathbf{u}_r$ ONB $R(A)$ & $d)\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ ONB $\mathcal{N}(A^T)$

7) $\text{rank}(A) = r < n$ we set $U_1 = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$ $V_1 = (\mathbf{v}_1, \ldots, \mathbf{v}_r)$ and

$$A = U_1 \Sigma_1 V_1^T \quad \text{called COMPACT FORM OF THE SVD OF } A \tag{6.5.6}$$

Proofs: in the book.

Application:

rank $(A) = n$ where $A \in \mathbb{R}^{n\times n}$ so A injective.
Claim: There exists A^T s.t. $A^T A = I_{n\times n}$
PROOF:

$$A = U\Sigma V^T \text{ with } \Sigma = \begin{pmatrix} \sigma_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \sigma_n \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \end{pmatrix}$$

$$\text{define}, A^T = V \left(\Sigma_1^{-1} \; 0 \right) U^T \Rightarrow A^T A = V \left(\Sigma_1^{-1} \; 0 \right) U^T U \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} V^T = VV^T = I_{n\times n}$$

LEFT INVERSE OF A is A^T. Is not unique: $A^T := V \left(\Sigma_1^{-1} \; R \right) U^T$ where $R \in \mathbb{R}^{n\times(m-n)}$ is arbitrary but besides that also satisfies $A^T A = I$

Note:

All above was for $n \leq m$ so tall and square case
What about the fat case so $A \in \mathbb{R}^{m\times n}$ where $m \leq n$
Solution: SVD of $A^T \in \mathbb{R}^{n\times m}$

$$A^T = \overline{U} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \overline{V}^T \quad \text{where } \Sigma_1 = \begin{pmatrix} \overline{\sigma_1} & & \\ & \ddots & \\ & & \overline{\sigma}_r \end{pmatrix}$$

$$\text{where } \overline{\sigma}_1 \geq \ldots \geq \overline{\sigma}_r > 0 \text{ and } r = \text{rank}(A^T)$$

$$\Rightarrow A = \overline{V} \begin{pmatrix} \Sigma_1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \overline{U}^T$$

$$\text{Define } U := \overline{V}, V := \overline{U} \Rightarrow \sigma_i = \overline{\sigma}$$

Consequence: do not distinguish them. One theorem:
Theorem:
$A \in \mathbb{R}^{m\times n}$ and rank $(A) = r \leq \min(n, m)$. Exists $\sigma_1 \geq \ldots \geq \sigma_r > 0$ and orthogonal $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ s.t. $A = U \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^T$
The southeast corner is equal to $(m - r) \times (n - r)$

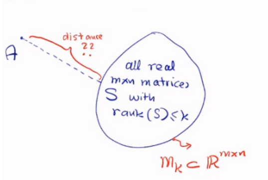
Note:

rank$(A) = n$ so injective then the O-matrices right side absent.
If rank$(A) = m$ so surjective then the O-matrices on bottom absent.
If A square and nonsingular, then all O-matrices absent.

Lecture 9:

Applications:

Let $A \in \mathbb{R}^{m\times n}$ and $\text{rank}(A) = r$ pick integer k with $0 \leq k < r$
How far is A away from having rank at most k



Define:

$$\text{Frobenius norm } \|M\|_F := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}^2} = \sqrt{\text{tr}(M^TM)}$$

$$M_k := \{S \in \mathbb{R}^{m\times n} | \text{rank}(S) \leq k\}$$

$$\text{Distance } d(A, M_k) = \inf\{\|A - S\|_F | S \in M_k\} \text{ where } \inf = \min$$

So we want to find $X \in M)k$ s.t. $\|A - X\|_F = d(A, M_k)$
Note: $M_k \subset \mathbb{R}^{m\times n}$ not a linear subspace so least squares not possible.

Some theorems:

Lemma 6.5.2
,and A is $m \times n$ and Q is $m \times m$ orthogonal, then $\|QA\|_F = \|A\|_F$
Proof:

$$\|QA\|_f^2 = \|Q\mathbf{a}_1, \ldots, Q\mathbf{a}_n\|_F^2 = \sum_{i=1}^{n} \|Q\mathbf{a}_i\|_2^2 = \sum_{i=1}^{n} \|\mathbf{a}_i\|_2^2 = \|A\|_F^2$$

because $\|A\|_F = \left\|\Sigma V^T\right\|_F$ it follows that

$$\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_n^2}$$

Theorem 6.5.3:
$A = U\Sigma V^T$ is $m \times n$ and M_k denotes the set of rank r or less where $0 < k < \text{rank}(A)$. If $X \in M_k$ then

$$\|A - X\|_F = \sqrt{\sigma_{k+1}^2 + \ldots + \sigma_n^2}$$

Proof:

$$\|A - X\|_F = \left\|U\Sigma V^T - X\right\|_F = \left\|\begin{pmatrix} \sigma_{k+1} & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}\right\|_F = \sqrt{\sigma_{k+1}^2 + \ldots + \sigma_n^2}$$

Special case of application:

Let $A \in \mathbb{R}^{n\times n}$ nonsingular, how far from singular.
$S \in \mathbb{R}^{n\times n}$ singular $\Leftrightarrow$ rank$(S) \leq n-1$. so distance of A being singular is $d(A, M_{n-1})$
Let $X \in M_{n-1}$ and $\|A - X\|_F = d(A, M_{n-1})$
$d(A, M_k) = \sqrt{(\sigma_{k+1}^2 + \sigma_{k+2}^2 + \ldots + \sigma_n^2)}$
After SVD of A and define $X := U \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^T$ where $\Sigma_1 = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$
We see that $X \in M_k$ and $\|A - X\|_F = d(A, M_k)$ so X is a best approximation in M_k of A
Other way:
Distance to singularity:
Note that $M_{n-1} \subset \mathbb{R}^{n\times n}$ is the set of all singular $n \times n$ matrix.
Accoarding theorem 6.5.3 we see that $d(A, M_{n-1}) = \|A - X\|_F = \sqrt{\sigma_n^2} = \sigma_n$
We have that $A = U\Sigma V^T$ with $\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}$ so then $A' = U\Sigma' V^T$ with $\Sigma' = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{n-1} \end{pmatrix}$ best approximation. If $A = U_1\Sigma_1 V_1^T$ we see that we can rewrite $A = \sum_{i=1}^{r} \sigma_i y_i v_i^T$ where $r = \text{rank}(A)$

Lecture 10:

Quadratic forms:

Definition 6.6.1:
QUADRATIC EQUATION two variables

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$
$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = 0 \tag{6.6.1,2}$$

And when we let $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$
Quadratic form associated with quadratic equation: $\mathbf{x}^T A\mathbf{x} = ax^2 + 2bxy + cy^2$

Theorem 6.6.1: PRINCIPAL AXES THEOREM:
$A \in \mathbb{R}^{n\times n}$ symmetric then change of variables $\mathbf{u} = Q^T\mathbf{x}$ s.t. $\mathbf{x}^T A\mathbf{x} = \mathbf{u}^T D\mathbf{u}$ where D diagonal.
PROOF:
A real symmetric, so by Corollary 6.4.7 we know that exists orthogonal Q that diagonalize A so $Q^T AQ = D$. When we substitute $\mathbf{u} = Q^T\mathbf{x}$ then $\mathbf{x} = Q\mathbf{u}$ and by substitution it is correct.

Definition 6.6.2: CLASSICAL CALCULUS PROBLEM
Let $F(\mathbf{x})$ real-valued function on $\mathbb{R}^n$. A point $\mathbf{x}_0 \in \mathbb{R}^n$,STATIONARY POINT OF F if all first partial derivatives of F exists at $\mathbf{x}_0$ and are zero.

Definition 6.6.4
Let $A \in \mathbb{R}^{n\times n}$ be symmetric.
(1) A POSITIVE DEFINITE if $\mathbf{x}^T A\mathbf{x} > 0$ then $\mathbf{x} \neq 0$ Notation: $A > 0$
(2) A NEGATIVE DEFINITE if $\mathbf{x}^T A\mathbf{x} < 0$ then $\mathbf{x} \neq 0$ Notation: $A < 0$
(3) A POSITIVE SEMI-DEFINITE if $\mathbf{x}^T A\mathbf{x} \geq 0$ then $\forall \mathbf{x} \in \mathbb{R}^n$ Notation: $A \geq 0$
(4) A NEGATIVE SEMI-DEFINITE if $\mathbf{x}^T A\mathbf{x} \leq 0$ then $\forall \mathbf{x} \in \mathbb{R}^n$ Notation: $A \leq 0$
(5) $A < 0 \Leftrightarrow -A > 0$ and $A \leq 0 \Leftrightarrow -A \geq 0$
(6) A INDEFINITE if $\mathbf{x}^T A\mathbf{x}$ cantake positive as negative real values.

Theorem 6.6.2:
$A \in \mathbb{R}^{n\times n}$ symmetric, $\lambda_i \in \mathbb{R}$ for $i = 1, \ldots, n$ eigenvalues. Then we have that

$$\begin{aligned}
&a) && A > 0 \Leftrightarrow \lambda_i > 0 \\
&\text{PROOF} \Rightarrow && \mathbf{x} \text{ corresponding eigenvector } \mathbf{x}^T A\&x = \lambda\mathbf{x}^T\mathbf{x} = \lambda\|\mathbf{x}\|^2 \\
& && \lambda = \frac{\mathbf{x}^T A\mathbf{x}}{\|\mathbf{x}\|^2} \text{ because both greater then } 0 \Rightarrow \lambda > 0 \\
&\text{PROOF} \Leftarrow && \{\mathbf{u}_1, \ldots, \mathbf{u}_n\} \text{ Orthonormal set eigenvectors } A \\
& && \mathbf{x} \in \mathbb{R}^n \Rightarrow \mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n \text{ where } c_i = \mathbf{x}^T\mathbf{u}_i \Rightarrow \mathbf{x}^T A\mathbf{x} = \sum_{i=1}^n c_i^2\lambda_i \\
& && \text{because } c_i \in \mathbb{R} \Rightarrow c_i^2 > 0 \text{ and } \lambda_i > 0 \Rightarrow \mathbf{x}^T A\mathbf{x} > 0 \Rightarrow A > 0 \\
&b) && A < 0 \Leftrightarrow \lambda_i < 0 \\
&c) && A \geq 0 \Leftrightarrow \lambda_i \geq 0 \\
&d) && A \leq 0 \Leftrightarrow \lambda_i \leq 0 \\
&e) && A \text{ indefinite} \Leftrightarrow \text{eigenvalues different sign}
\end{aligned}$$

Hessian matrix:

General case:

For C^2 function $\mathbb{R}^2 \to \mathbb{R}$ and stationary point $\mathbf{x}_0 \in \mathbb{R}^n$ Hessian matrix:

$$H(\mathbf{x}_0) = (h_{ij}) \text{ where } h_{ij} = F_{x_i,x_j}(\mathbf{x}_0)$$

3 options:
(1) $H(\mathbf{x}_0) > 0 \Rightarrow \mathbf{x}_0$ local minimum
(2) $H(\mathbf{x}_0) < 0 \Rightarrow \mathbf{x}_0$ local maximum
(3) $H(\mathbf{x}_0)$ indefinite $\Rightarrow \mathbf{x}_0$ saddle point.

Leading principal submatrix of order r: the upperleft corner A_r of size $r \times r$ in the main square matrix.
Example:
$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 4 \end{pmatrix}$ then $A_1 = 1$ and $A_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and $A_3 = A$

Theorem 6.7(.1):

$A \in \mathbb{R}^{n\times n}$ symmetric then following properties:

Property 1 $A > 0 \Leftrightarrow \det(A) > 0$

proof $\det(A) = \prod_{i=1}^{n} \lambda_i > 0$

Property 2 $A > 0 \Leftrightarrow A$ nonsingular

Proof by definition $\det(A) \neq 0 \Leftrightarrow A$ nonsingular

$A \in \mathbb{R}^{n\times n}$ then following 5 equivalent:

1) $A > 0$
2) $\det(A_r) > 0$ for $r = 1, \dots, n$
3) A can be reduced to upper triangular by only row operation 3
4) $\exists$lower triangular L positive elements s.t. $A = LL^T$
5) $\exists$nonsingular B s.t. $A = B^T B$

Proof:

$5 \Rightarrow 1$ $A = B^T B$ and $\mathbf{x} \neq \mathbf{0} \Rightarrow \mathbf{x}^T A\mathbf{x} = \|B\mathbf{x}\|^2 > 0 \Rightarrow A > 0$

$1 \Rightarrow 2$ $A > 0$ take $1 \leq r \leq n \Rightarrow A = \begin{pmatrix} A_r & * \\ * & * \end{pmatrix}$ .let $\mathbf{x} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$ where $x_1 \neq 0$

$\Rightarrow \mathbf{x}^T A\mathbf{x} = \mathbf{x}_1^T A_r \mathbf{x}_1 \Rightarrow$ left part greater then 0, so right part so $\det(A_r) > 0$

$2 \Rightarrow 3$ See Book

$3 \Rightarrow 4$ See Book

$4 \Rightarrow 5$ $A = LL^T \Rightarrow B = L^T \Rightarrow A = B^T(B^T)^T = B^T B$

Lecture 11:

$A \in \mathbb{C}^{n\times n}$ with $p_i \in \mathbb{C}$ for $i = 1, \ldots, n$ then CHARACTERISTIC POLYNOMIAL:

How we learned it: $p_A(t) = \det(A - tI) = (-1)^n(t^n + p_{n-1}t^{n-1} + \ldots + p_1 t + p_0)$

More general: $p_A(t) = p_n t^n + p_{n-1}t^{n-1} + \ldots + p_1 t + p_0$

So then we have that: $p_A(A) = p_n A^m + p_{n-1}A^{n-1} + \ldots + p_1 A + p_0 I$

Cayley Hamilton theorem: $p_A(A) = O$ where $O \in \mathbb{R}^{n\times n}$ is the zero-matrix.
Example:
$A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ then $\rho_A(t) = t^2$ and $\rho_A(A) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)^2 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$

Proof linearly dependent:

$\mathbb{C}^{n]timesn}$ has dimension n^2. Hence $n^2 + 1$ matrices $I, A, A^2, \ldots, A^{n^2}$ lin. dependent.
So $\exists \alpha_0, \alpha_1, \ldots, \alpha_{n^2}$ s.t. $\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \ldots + \alpha_{n^2} A^{n^2} = 0$

Accoarding the Cayley Hamilton theorem:
1) Already $I, A, A^2, \ldots, A^n$ are linearly dependent.
2) Required coeeficients are in characteristic polynomial A:

$$p_A(t) = (-1)^n(t^n + p_{n-1}t^{n-1} + \ldots + p_1 t + p_0) \Rightarrow p_0 I + p_1 A + \ldots + p_{n-1}A^{n-1} + A^n = 0$$

2 Lemma's:

Let $A, B \in \mathbb{R}^{n\times n}$

Lemma 1: $\exists$ nonsingular P s.t. $B = P^{-1}AP$, then $A\&B$ same characteristic polynomial

PROOF $B = P^{-1}AP$ so B similar to $A \Rightarrow p_A(s) = p_B(s)$

Lemma 2: $B = \begin{pmatrix} 0 & 0 & \ldots & 0 & -b_0 \\ 1 & 0 & \ldots & 0 & -b_1 \\ 0 & 1 & \ldots & 0 & -b_2 \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & \ldots & 1 & -b_{n-1} \end{pmatrix} \Rightarrow p_B(t) = (-1)^n(t^n + b_{n-1}t^{n-1} + \ldots + b_1 t + b_0)$

PROOF $n = 2 : B = \left(\begin{smallmatrix} 0 & -b \\ 1 & -b_1 \end{smallmatrix}\right) \Rightarrow p_B(t) = \left|\begin{smallmatrix} -t & -b_0 \\ 1 & -b_1 - t \end{smallmatrix}\right| = t^2 + b_1 t + b_0$

so statement is true for $n = 2$

assume true for $n - 1 \Rightarrow p_A(t) = (-1)^{n-1}(t^{n-1} + b_{n-2}t^{n-2} + \ldots + b_1 t + b_0)$

now for n $\det(B - tI) = \begin{vmatrix} -t & 0 & \ldots & 0 & -b_0 \\ 1 & -t & \ldots & 0 & -b_1 \\ 0 & 1 & \ldots & 0 & -b_2 \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & \ldots & 1 & -b_{n-1} - t \end{vmatrix} = -t \cdot p_A + (-1)^n b_0$

when you work this out you get indeed that

$p_B(t) = (-1)^n(t^n + b_{n-1}t^{n-1} + \ldots + b_1 t + b_0)$

Proof of cayley hamilton:

To prove $A^n + p_{n-1}A^{n-1} + \ldots + p_1 A + p_0 I = O_{n\times n}$

$\boxed{A}$ take arbitrarily $1 \leq k \leq n$ s.t. u_i is the k th column of A^i
So then $u_n + p_{n-1}u_{n-1} + \ldots + p_1 u_1 + p_0 u_0 = 0$
assume $u_0, \ldots, u_{n-1}$ lin. independent,so $u_n + b_{n-1}u_{n-1} + \ldots + b_1 u_1 + b_0 u_0 = 0$
now we want to prove that $b_i = p_i$ for $i = 0, \ldots, n-1$
note that $u_0 = e_k, u_1 = Au_0, \ldots, u_n = Au_{n-1}$

$\boxed{B}$ Matrix B of lin. map A w.r.t. basis $\{u_0, \ldots, u_{n-1}\}$
has characteristic polynomial $p_B(t) = (-1)^n(t^n + b_{n-1}t^{n-1} + \ldots + b_1 t + b_0)$

Proof of $\boxed{B}$
$$\begin{cases} Au_0 = u_1 = 0\cdot u_1 + 1\cdot u_1 + \ldots + 0\cdot u_{n-1} \\ \vdots \\ Au_{n-2} = 0\cdot u_0 + \ldots + 1\cdot u_{n-1} \\ Au_{n-1} = -b_0u_0 - b_1u_1 + \ldots - b_{n-1}u_{n-1} \end{cases} \Rightarrow \text{matrix similair to Lemma 2}$$

By Lemma 2 $P_B(t) = (-1)^n(t^n + b_{n-1}t^{n-1} + \ldots + b_1 t + b_0)$

$\boxed{C}$ $A\&B$ similair so $P_A(t) = P_B(t)$ So therefore $b_i = p_i$ for $i = 0, \ldots, n-1$
so $u_n + p_{n-1}u_{n-1} + p_1 u_1 + p_0 u_0 = 0$ so under assumption that
$u_0, \ldots, u_{n-1}$ lin. independent and k arbitary,we proved the CH

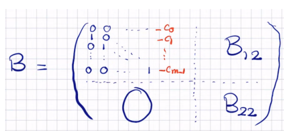
Proof General case:

$\boxed{A}$ Define u for arbitrarily k as follows $u_i :=$ k th column of A^i for $i = 0, \ldots, n$
let $m < n$ s.t. $u_0, \ldots, u_{m-1}$ lin. independent,but $u_0, \ldots, u_m$ not
so there exists $c_0, \ldots, c_{m-1}$ s.t. $u_m + c_{m-1}u_{m-1} + \ldots + c_0 u_0 = 0$
Define $q(t) = t^m + c_{m-1}t^{m-1} + \ldots + c_1 t + c_0$

$\boxed{B}$ Claim $p_A(t)$ divisible by $q(t)$ so for polynomial $r(t) \Rightarrow p(t) = q(t)r(t)$

Proof of $\boxed{B}$ extend $\{u_0, \ldots, u_{m-1}\}$ to arbitrarily basis $\{u_0, \ldots, u_{m-1}, v_m, \ldots, v_{n-1}\}$ of $\mathbb{C}$

$$\begin{cases} Au_0 = u_1 = 0\cdot u_0 + 1\cdot u_1 + \ldots \\ Au_1 = u_2 = 0\cdot u_0 + 0\cdot u_1 + 1u_2 + \ldots \\ \vdots \\ Au_{m-1} = u_m = -c_0u_0 - c_1u_1 + \ldots - c_{m-1}u_{m-1} \\ Av_m = \ldots \\ \vdots \\ Av_{n-1} = \ldots \end{cases} \Rightarrow$$



$P_B(t) = \det(B - tI) = q(t)\det(B_{22} - tI)$ since $P_B(t) = p_A(t) \Rightarrow p_A(T) = p_B(t) = q(t)r(t)$

$\boxed{c}$ for any $1 \leq k \leq n \Rightarrow p_A(A)e_k = r(A)q(A)e_k$
$= r(A)(A^m + \ldots + c_1 A + c_0 I)e_k = r(A)(u_m + c_{m-1}u_{m-1} + \ldots + c_1 u_1 + c_0 u_0) = r(A)0 = 0$
$p_A(A) = 0$

Lecture 12:

Beginning of the simple form:

Theorem 1
$A \in \mathbb{C}^{n\times n}$ and $T \in \mathbb{C}^{n\times n}$ invertible then:

$$T^{-1}AT = J \quad J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix} \text{ where } J_i = \lambda I + N = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

We call here J JORDAN NORMAL FORM and J_i JORDAN BLOCK
To go further we have some tools:

Tools:

KERNEL/NULL SPACE $\ker(A) := \{\mathbf{x} \in V : A\mathbf{x} = \mathbf{0}\}$
RANGE $\text{range}\,(A) := \{A\mathbf{x} : \mathbf{x} \in V\}$
$\dim\ker(A) + \dim\text{range}(A) = n$
$\mathbf{x}$ eigenvector of A corresponding to eigenvalue A if $A\mathbf{x} = \lambda\mathbf{x}$
EIGENSPACE ASSOCIATED WITH λ : $\ker(A - \lambda I)$
GEOMETRIC MULTIPLICITY OF λ : $\dim\ker(A - \lambda I)$
λ eigenvlaue iff root of characteristic polynomial: $p_z = \det(A - zI)$
$p(z) = (-1)^n (z-\lambda_1)^{a_1}(z-\lambda_2)^{a_2}\dots(z-\lambda_k)^{a_k}$
where $\lambda_1, \dots, \lambda_k$ distinct eigenvalues A and a_j ALGEBRAIC MULTIPLICITY.
corresponding geometric multiplicity g_j

Decomposition invariant subspaces:

$V_1, \dots, V_k \subset V$. V DIRECT SUM of $V_1, \dots, V_k$ if each $\mathbf{x} \in V$ can be written unique:

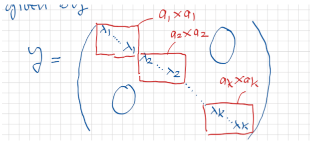
$$\begin{aligned} \mathbf{x} &= \mathbf{x}_1 + \dots + \mathbf{x}_k \quad \text{where} \quad \mathbf{x}_j \in V_j \quad j = 1, \dots, k \\ \overline{\mathbf{x}} &= \overline{\mathbf{x}_1} + \dots + \overline{\mathbf{x}_k} \quad \text{with } \overline{\mathbf{x}_i} \in V_i \\ &\Rightarrow \mathbf{x}_i = \overline{\mathbf{x}_i} \quad \text{for } i = 1, \dots, k \\ \text{notation} V &= V_1 \bigoplus V_2 \bigoplus \cdots \bigoplus V_k \\ W \subset V & \text{ INVARIANT UNDER } A: \ \mathbf{x} \in W \Rightarrow A\mathbf{x} \in W \end{aligned}$$

Proof special case:

situation $a_i = g_i \quad i = 1, \ldots, k \quad \lambda_1, \ldots, \lambda_k$ distinct eigenvalues

To prove: A has n lin. indep. eigenvectors

A diagonalizable



$J_i \in \mathbb{C}^{1\times 1}$, these blocks are λ_i, (a_i times) where $i \in \{1, k\}$

A Construct basis $\mathbb{C}^n$ s.t. J matrix of A w.r.t. that basis

by assumtion: $\dim \mathcal{N}(A - \lambda_i I) = g_i = a_i$

Choose basis $\{u_1^i, \ldots, u_{a_i}^i\}$ of $N(A - \lambda_i I)$

$a_1 + \ldots + a_k = n \Rightarrow n$ vectors in $\mathbb{C}^n$

B claim these vectors linearly independent, basis of $\mathbb{C}^n$

SUBPROOF: $\mathcal{N}(A - \lambda_j I) \ni \mathbf{v}_i = \alpha_1^i u_1^i + \alpha_2^i u_2^i + \ldots + \alpha_{a_i}^i u_{a_i}^i$

$$\sum_{i=1}^{k} \mathbf{v}_i = 0 \Leftrightarrow \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_k = 0$$

I will show that $\mathbf{v}_i = 0 \quad (i = 1, \ldots, k)$.assume not

$v_1, v_2, \ldots, v_r \neq 0 \quad \& \quad v_{r+1}, \ldots, v_k = 0$

Note: $i = 1, \ldots, r$, the vector $\mathbf{v}_i$ eigenvector with eigenvector λ_i

since $\lambda_1, \ldots, \lambda_r$ distinct, the corresponding vectors are linearly independent

however $v_1 + \ldots + v_r = 0$ this is impossible, so contradiction

C the vectors $\mathbf{u}_j^i \quad i = 1, \ldots, k \& j = 1, \ldots, a_i$ basis $\mathbb{C}^n$

$$\mathbb{C}^n = \mathcal{N}(A - \lambda_1 I) \bigoplus \ldots \bigoplus \mathcal{N}(A - \lambda_k I)$$

Note: each $\mathcal{N}(A - \lambda_i I)$ invariant

SUBPROOF $\mathbf{x} \in \mathcal{N}(A - \lambda_i I) \Rightarrow A\mathbf{x} = \lambda_i \mathbf{x} \Rightarrow AA\mathbf{x} = \lambda_i A\mathbf{x} \Rightarrow (A - \lambda_i I)A\mathbf{x} = 0 \Rightarrow A\mathbf{x} \in \mathcal{N}(A - \lambda_i I)$

D Diagonal matrix J of A w.r.t $\{\mathbf{u}_j^i | i = 1, \ldots, k \& j = 1, \ldots, a_i\}$

when we write this out for $i = 1 \& j = 1, \ldots, a_j$ we get $\begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_1 & 0 & \ldots & 0 \\ \vdots & \vdots & & & \\ 0 & 0 & \ldots & 0 & \lambda_i \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$

total number of rows with a_1 is equal to a_1

when we continue this, we find indeed the Jordan canonical form we had to obtain

Equivalent statements: for $i = 1, \ldots, k$

1) $a_i = g_i$ 2) $\dim(A - \lambda_i I) = a_i$ 3)Jordan blocks like above 4)Jordan form is diagonal 5) A diagonalizable

Lecture 13:

situation $\mathbb{C}^n = \mathcal{N}(A-\lambda_1 I)^{m_1} \bigoplus \mathcal{N}(A-\lambda_2 I)^{m_2} \bigoplus \cdots \bigoplus \mathcal{N}(A-\lambda_k I)^{m_k} \quad m_i \in \mathbb{Z}^+$
$\mathcal{N}(A-\lambda_i I)^{m_i}$

Strategy : Choose for each V_i a basis $\{\mathbf{u}_1^i, \ldots, \mathbf{u}_{a_i}^i\} \Rightarrow$ all together basis $\mathbb{C}^n$

since $AV_i \subset V_i \Rightarrow B$ of A w.r.t. basis is blockdiagonal $B = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$

find correct bases for the matrices A_i

$A \in \mathbb{C}^{n\times n}$ with characteristic polynomial $p_A(z)$ where $\deg(p_A(z)) = n$
Definition:
$q(z)$ polynomial. $q(z)$ ANNIHILATES A if $q(A) = 0$
Example: By CH-Theorem $p(z)$ annihilates A.

Theorem:
$A \in \mathbb{C}^{n\times n} \Rightarrow \exists$ one monic polynomial $p_{\min}(z)$ of minimal degree tha annihilates A.
PROOF:

$\boxed{A}$ $S_A := \{$set of polynomials annihilate $A\}$, take polynomial smallest degree: $p_{\min}(z)$
$\tilde{p}_{\min}(z) \in S_A \Rightarrow \deg(p_{\min}(z)) = \deg(\tilde{p}_{\min}(z)) \Rightarrow p_{\min}(z) = \tilde{p}_{\min}(z)$

$\boxed{\text{Lemma}}$ $p(z)$ any polynomial annihilates $A \Rightarrow \exists q(z)$ (polynomial) s.t. $p(z) = p_{\min}(z)q(z)$

PROOF LEMMA Division algorithm: $p(z) = q(z)p_{\min}(z) + r(z)$

$\boxed{\text{option 1}}$ $r(z) \neq 0$ and $\deg(r(z)) < \deg(p_{\min}(z))$
$r(A) = p(A) - q(A)p_{\min}(A) = 0 - q(A)0 = 0 \Rightarrow r(z)$ annihilates A
and $\deg(r(A)) < \deg(p_{\min}(z)) \Rightarrow$ contradiction

$\boxed{\text{option 2}}$ $r(z) = 0 \Rightarrow p(z) = q(z)p_{\min}(z)$

$\boxed{\text{Continue with } A}$ $p_{\min}(z)$ monic polynomial min. degree annihilates A.
$\tilde{p}_{\min}(z)$ second monic polynomial min. degree annihilates A.
$\tilde{p}_{\min}(z) = q(z)p_{\min}(z)$ because $\deg(p_{\min}(z)) = \deg(\tilde{p}_{\min}(z)) \Rightarrow q(z) = q_0 \in \mathbb{R}$
$\tilde{p}_{\min}(z) \& p_{\min}(z)$ monic so $q_0 = 1 \Rightarrow p_{\min}(z) = \tilde{p}_{\min}(z) \Rightarrow p_{\min}(z)$ unique

Theorem:
$A \in \mathbb{C}^{n\times n}$ every eigenvalue of A root of $p_{\min}(z)$ converse also true.
PROOF:

$\boxed{A}$ $p_A(z)$ annihilates $A \Rightarrow p_A(z) = q(z)p_{\min}(z)$ for some $q(z)$
$p_{\min}(\lambda) = 0 \Rightarrow p_A(\lambda) = 0 \Rightarrow \lambda$ eigenvalue A with corresponding $\mathbf{x}$ s.t. $A\mathbf{x} = \lambda\mathbf{x}$

$\boxed{\text{claim}}$ for any $p(z) \Rightarrow p(A)\mathbf{x} = p(\lambda)\mathbf{x}$

PROOF $p(z) = p_k z^k + \ldots + p_1 z + p_0 \Rightarrow p(A) = p_k A^k + p_{k-1}A^{k-1} + \ldots + p_1 A + p_0 I$

$\boxed{\text{back to } A}$ $p_{\min}(z) \Rightarrow p_{\min}(A)\mathbf{x} = p_{\min}(\lambda)\mathbf{x}$ since $p_{\min}(A) = 0 \& \mathbf{x} \neq \mathbf{0} \Rightarrow p_{\min}(\lambda) = 0$
$\Rightarrow$ roots of $p_{\min}(z) \Rightarrow p_{\min}(z) = (z-\lambda_1)^{m_1} \ldots (z-\lambda_k)^{m_k}$

Theorem:
$a \in \mathbb{C}^{n\times n}$ and λ_i for $i = 1, \dots, k$ distinct eigenvalues, m_i for $i = 1, \dots, k$ integers like above. Then $V_i = \mathcal{N}((A - \lambda_i I)^{m_i})$ satisifies:

$$1) V_i \text{ is } A\text{-invariant} \quad 2) \mathbb{C}^n = V_1 \bigoplus \cdots \bigoplus V_k$$

For the proof we need the following
Lemma 2:
$A \in \mathbb{C}^{n\times n}$ let $p(z)$ polynomial s.t. $p(A) = 0$. If $p(z) = p_1(z)p_2(z)$ where those two polynomials have no common root(so COPRIME),then:

$$\mathcal{N}(p_1(A)) \quad \text{and} \quad \mathcal{N}(p_2(A)) \text{ are } A \text{ invariant}$$
$$\mathbb{C}^n = \mathcal{N}(p_1(A)) \bigoplus \mathcal{N}(p_2(A))$$

PROOF:

$\boxed{A}$ $p_1(z)q_1(z) + p_2(z)q_2(z) = 1 \Rightarrow$ BÉZOUT IDENTITY $\Rightarrow p_1(A)q_1(A) + p_2(A)q_2(A) = I$
$\mathbf{x} \in \mathbb{C}^n \Rightarrow \mathbf{x} = p_1(A)q_1(A)\mathbf{x} + p_2(A)q_2(A)\mathbf{x} = \mathbf{x}_2 + \mathbf{x}_1$

$\boxed{\text{Claim}}$ $\mathbf{x}_2 \in \mathcal{N}(p_2(A))$

PROOF $p_2(A)\mathbf{x} = p_2(A)p_1(A)q_1(A)\mathbf{x} = p(A)q_1(A)\mathbf{x} = q_1(A)\mathbf{x} = 0$
likewise $\mathbf{x}_1 \in \mathcal{N}(p_1(A))$

$\boxed{\text{back to } A}$ so $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \Rightarrow \mathbb{C}^n = \mathcal{N}(p_1(A)) + \mathcal{N}(p_2(A))$

$\boxed{\text{proof of the direct sum}}$ suppose $\mathbf{x} = \mathbf{x}_1' + \mathbf{x}_2'$ with $\mathbf{x}_i' \in \mathcal{N}(p_i(A))$
$\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_1' + \mathbf{x}_2' \Leftrightarrow \mathbf{x}_1 - \mathbf{x}_1' = \mathbf{x}_2' - \mathbf{x}_2$
call this vector $\mathbf{u} : \mathbf{u} = (q_1(A)p_1(A) + q_2(A)p_2(A))\mathbf{u} = 0$
since $p_1(A)\&u = 0\&p_2(A)\mathbf{u} = 0 \Rightarrow \mathbf{u} = 0 \Rightarrow \mathbf{x}_1 = \mathbf{x}_1'\&\mathbf{x}_2 = \mathbf{x}_2'$
unique decomposition $\mathbb{C}^n = \mathcal{N}(p_1(A)) \bigoplus \mathcal{N}(p_2(A))$
statement 2 of the Lemma is proven
now statement 2
$\mathbf{x} \in \mathcal{N}(p_1(A)) \Rightarrow p_1(A)\mathbf{x} = 0 \Rightarrow Ap_1(A)\mathbf{x} = 0 \Rightarrow Ap_1(A) = p_1(A)A$
$\Rightarrow p_1(A)A\mathbf{x} = 0 \Rightarrow A\mathbf{x} \in \mathcal{N}(p_1(A))$
simlair for $\mathcal{N}(p_2(A))$

PROOF OF THE THEOREM:

$p_{\min}(z) = (z - \lambda_1)^{m_1} \dots (z - \lambda_k)^{m_k} \Rightarrow (z - \lambda_i)^{m_i}$ coprime since $\lambda_1, \dots, \lambda_k$ distinct
by Lemma $\mathbb{C}^n = V_1 \bigoplus \cdots \bigoplus V_k$ with $V_i = \mathcal{N}((A - \lambda_i I)^{m_i})\, quad(i = 1, \dots, k)$
each V_i is A-invariant

Lecture 14:

So we have $A|_{V_i} : V_i \to V_i$ where $V_i = N((A-\lambda_i I)^{m_i})$

Lemmas:

Lemma 1 of this lecture:
$A|_{V_i}$ 1 eigenvalue λ_i
PROOF:

$$\boxed{A} \quad \exists \mathbf{x} \in V_i\, \mathbf{x} \neq \mathbf{0} \text{ s.t. } A|_{V_i}\mathbf{x} = \lambda\mathbf{x} \Rightarrow A\mathbf{x} = \lambda\mathbf{x} \Rightarrow (A-\lambda_i I)^{m_i}\mathbf{X} = 0$$

$$\boxed{B} \quad \text{Claim } (A-\lambda_i I)^{m_i}\mathbf{x} = (\lambda-\lambda_i)^{m_i}\mathbf{x}$$

$$\boxed{\text{PROOF } B} \quad (A-\lambda_i I)\mathbf{x} = A\mathbf{x} - \lambda_i\mathbf{x} = (\lambda-\lambda_i)\mathbf{x}$$

$$\xrightarrow{\text{by induction you can show that}} (A-\lambda_i I)^{m_i}\mathbf{x} = (\lambda-\lambda_i)^{m_i}\mathbf{x}$$

$$\boxed{\text{back to } A} \quad (\lambda-\lambda_i)^{m_i}\mathbf{x} = 0 \,\&\, \mathbf{x} \neq 0 \Rightarrow \lambda = \lambda_i$$

Lemma 2 of this lecture:
$\dim(V_i) = a_i$ algebraic multiplicity of λ_i
PROOF:

$$\mathbb{C}^n = V_1 \bigoplus \cdots \bigoplus V_n \Rightarrow A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$$

$$\Rightarrow p_A(z) = \det(A-zI) = \begin{vmatrix} A_1 - zI & & \\ & \ddots & \\ & & A_k - zI \end{vmatrix} = \det(A_1 - zI)\ldots\det(A_k - zI)$$

$$\xrightarrow{\text{Lemma 1}} A_i \text{ has 1 eigenvalue, for now } \dim(V_i) = n_i \text{ Still unknown}$$

$$\det(A_i - zI) = (-1)^{n_i}(z-\lambda_i)^{n_i} \text{ so } n_1 + \ldots + n_k = n$$

$$p_A(z) = (-1)^n (z-\lambda_1)^{n_1}\ldots(z-\lambda_k)^{n_k} \Rightarrow n_i = a_i \quad \text{for } i = 1,\ldots,k$$

Lemma 4 of this lecture:
Geometric multiplicity of eigenvalue of $A|v_i$ denoted by λ_i is equal to g_i
PROOF:

$$\boxed{A} \quad \text{Geometric multiplicty} = \dim(\mathcal{N}(A|V_i - \lambda_i I) \subset V_i$$

$$\boxed{B} \quad \text{Claim: } N(A|V_i - \lambda_i I) = N(A-\lambda_i I) \cap V_i$$

$$\boxed{\text{PROOF} \Rightarrow} \quad \mathbf{x} \in \mathcal{N}(A|V_i - \lambda_i I) \Rightarrow \mathbf{x} \in V_i \,\&\, A|V_i\mathbf{x} = \lambda_i\mathbf{x} \Rightarrow A\mathbf{x} = \lambda_i\mathbf{x} \Rightarrow \mathbf{x} \in \mathcal{N}(A-\lambda_i I) \cap V_i$$

$$\boxed{\text{PROOF} \Leftarrow} \quad \mathbf{x} \in \mathcal{N}(A-\lambda_i I) \cap V_i \Rightarrow A|V_i\mathbf{x} = A\mathbf{x} = \lambda_i\mathbf{x} \Rightarrow \mathbf{x} \in \mathcal{N}(A|V_i - \lambda_i I)$$

$$\boxed{\text{continue with } A} \quad \dim(N(A-\lambda_i I) \cap V_i) : \text{note that } \mathcal{N}(A-\lambda_i I) \subset \mathcal{N}((A-\lambda_i I)^{m_i}) = V_i$$

$$\dim(\mathcal{N}(A-\lambda_i I) \cap V_i) = \dim \mathcal{N}(A-\lambda_i I) = g_i$$

Lemma 3 of this lecture:
Linear map: $A|V_i$ so the $a_i \times a_i$ matrix A_i has minimal polynomial $(z-\lambda_i)^{m_i}$
PROOF:

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & 0 & A_k \end{pmatrix}$$

definition minimal polynomial of $A:\ p_{\min}(z) = (z-\lambda_1)^{m_1} \dots (z-\lambda_k)^{m_k}$

for $i = 1, \dots, k: \begin{pmatrix} A_1-\lambda_i I & & & \\ & A_2-\lambda_i I & & \\ & & \ddots & \\ & & & A_k-\lambda_i I \end{pmatrix}^{m_i} = \begin{pmatrix} (A_1-\lambda_i I)^{m_i} & & & \\ & (A_2-\lambda_i I)^{m_i} & & \\ & & \ddots & \\ & & & (A_k-\lambda_i I)^{m_i} \end{pmatrix}$

$p_{\min}(A_i) = 0 \Leftrightarrow (A_i-\lambda_i I)^{m_1} \dots (A_i-\lambda_k I)^{m_k} = 0 \Rightarrow \boxed{1}$
by Lemma 1 one eigenvalue λ_i.when $j \neq i$ the matrices $A_i - \lambda_j I$ nonsingular $\Rightarrow (A_i-\lambda_j I)^{m_j}$ nonsingular
since $\boxed{1}$ we obtain $(A_i-\lambda_i I)^{m_i} = 0 \Rightarrow (z-\lambda_i)^{m_i}$ annihilates A_i
good candidate minimal polynomial A_i but assume that $(z-\lambda_i)^{l_i}$ annihilates A_i with $l_i < m_i$
$\Rightarrow (A_i-\lambda_i I)^{l_i} = 0 \Rightarrow$ define $q(z) = (z-\lambda_1)^{m_1} \dots (z-\lambda_i)^{l_i} \dots (z-\lambda_k)^{m_k}$
$q(A) = 0$, but $\deg(q(z)) < \deg(p_{\min}(z)) \Rightarrow$ contradiction,so $(z-\lambda_i)^{m_i}$ minimal polynomial

Single linear map: $A: \mathcal{V} \times \mathcal{V}$ with

1) $\dim(V) = A$ 2) eigenvalue$(A) = \lambda$ 3) $p_{\min}(z) = (z-\lambda)^m$ 4) geometric multiplicity $\lambda = g$

For map A we want to construct basis $\mathcal{V}$ s.t. A in Jordan form.
Special case:
Characteristic polynomial is minimal polynomial. We use notation

notation	$N: A-\lambda I \quad N^m = 0 \quad N^{m-1} \neq 0$
properties	$\exists \mathbf{u} \in \mathcal{V}, \mathbf{u} \neq 0$ s.t. $N^{m-1}\mathbf{u} \neq \mathbf{0} \quad N^m\mathbf{u} = \mathbf{0}$
A	Claim $\{N^{m-1}, \dots, N\mathbf{u}, \mathbf{u}\}$ basis $\mathcal{V}$
PROOF A	see lecture notes
B	Fact: A w.r.t. basis above of V is equal to the $m \times m$ matrix $\begin{pmatrix} \lambda & 1 & 0 & \dots & & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$
PROOF B	see lecture notes

Normal case:
JORDAN CHAIN set of nonzero vectors $\{N^{l-1}\mathbf{u}, \ldots, N\mathbf{u}, \mathbf{u}\}$ s.t. $N^l\mathbf{u} = 0$
LENGTH OF JORDAN CHAIN: $l \in \mathbb{Z}^+$ SINGLE VECTOR $\mathbf{u} \neq \mathbf{0}$ if $l = 1$

Fact:

$m \leq n \Rightarrow \mathcal{V}$ basis consting finitely many Jordain chains, so basis consisting of nonzero vectors
so $N^{l_i}\mathbf{u}_i = \mathbf{0} \quad \forall i \in \{1, \ldots, k\}$
$N\mathbf{u}_{k+j} = \mathbf{0} \quad \forall j \in \{1, \ldots, d\}$
Obviously $l_1 + \ldots + l_{k+d} = a$

PROOF: skipped.
Consequence:

$$\begin{cases} AN^{l_i-1}\mathbf{u}_i = \lambda N^{l_i-1}\mathbf{u}_i \\ AN^{l_i-2}\mathbf{u}_i = N^{l_i-1}\mathbf{u}_i + \lambda N^{l_i-2}\mathbf{u}_i \\ \vdots \\ AN\mathbf{u}_i = 1N^2\mathbf{u}_i + \lambda N\mathbf{u}_i \\ A\mathbf{u}_i = 1N\mathbf{u}_i + \lambda\mathbf{u}_i \end{cases} \quad \text{for } i = 1, \ldots, k$$

$$\begin{cases} A\mathbf{u}_{k+1} = \lambda\mathbf{u}_{k+1} \\ \vdots \\ A\mathbf{u}_{k+d} = \lambda\mathbf{u}_{k+d} \end{cases} \Rightarrow \text{See } \boxed{*}$$

$$\begin{cases} J \in \mathbb{R}^{a\times a} \text{ with } a = \dim(V), k+d \text{ Jordan blocks, d of them } 1\times 1. \\ k+d \text{ lin. independent } n^{l_i-1}\mathbf{u}_i \text{ for } i = 1, \ldots, k \& \mathbf{u}_{k+1}, \ldots, \mathbf{u}_{k+d} \\ \Rightarrow g = k+d \\ p_{\min}(A) = (z-\lambda)^m \text{ and } l_1 + \ldots + l_k = m \Rightarrow \deg(m) = \text{size largest Jordan Block} \\ \text{Jordan block not unique, order can be changed} \end{cases}$$

$\boxed{*}$