

# MAT 3341

## Applied Linear Algebra

### Study Guide



winter 2024

$$Ax = b \quad Ax = \lambda x$$

$$A = LU \quad PA = LU \quad A = QR$$

## Matrix Arithmetic

• addition / subtraction:

$$A = (a_{ij}) \quad B = (b_{ij}) \quad m \times n$$

$$A \pm B = (a_{ij} \pm b_{ij})$$

• scalar multiplication:

$$A = (a_{ij}) \quad \alpha \text{ is a scalar}$$

$$\alpha A = (\alpha a_{ij})$$

• matrix-vector multiplication:

$$A = (a_1, a_2, \dots, a_n)$$

$$a_1, a_2, \dots, a_n \in \mathbb{F}^m$$

$$x \in \mathbb{F}^n$$

$$Ax = \sum_{i=1}^n a_i x_i \in \mathbb{F}^m$$

• matrix-matrix multiplication

$$A: m \times q \text{ matrix}$$

$$B: q \times n \text{ matrix with columns } b_1, b_2, \dots, b_n \in \mathbb{F}^n$$

$$AB = (Ab_1, Ab_2, \dots, Ab_n)$$

or if  $C = AB$ , we have that

$$c_{ij} = \sum_{k=1}^q a_{ik} b_{kj}$$

## 2 special Matrices

$$\text{Zero matrix: } O_{m \times n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$$

$$\text{Identity matrix: } I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

For a  $m \times n$  matrix  $A$ :

$$A + O = O + A = A$$

$$I_m A = A I_n = A$$

## Diagonal Matrix

$$A = \text{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix} \in \mathbb{F}^{n \times n}$$

## Transpose

The transpose of a matrix  $A$  is a matrix  $A^T$  whose  $(i, j)$  entry is the  $(j, i)$  entry of  $A$ .

$$A = (a_{ij}) \in \mathbb{F}^{m \times n} \quad A^T = (a_{ji}) \in \mathbb{F}^{n \times m}$$

$A^T = A \Rightarrow A$  is symmetric

Some relations:

$$(A^T)^T = A$$

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

$$(\alpha A)^T = \alpha A^T$$

$$(AB)^T = B^T A^T$$

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

## Inverse

A matrix  $A \in \mathbb{F}^{n \times n}$  is said to be invertible if  $\exists X \in \mathbb{F}^{n \times n}$  s.t.  $XA = AX = I_n$

$X$  is an inverse of  $A$ , write  $A^{-1}$

Proposition: The inverse of an invertible matrix is unique.

If  $A$  and  $B$  are  $n \times n$  invertible matrices, then we have

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

## Link to Linear Map

A function  $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$  is a linear map if:

$$\cdot f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) \quad \forall \vec{x}, \vec{y} \in \mathbb{F}^n$$

$$\cdot f(\alpha \vec{x}) = \alpha f(\vec{x}) \quad \forall \alpha \in \mathbb{F}, \vec{x} \in \mathbb{F}^n$$

Proposition: Let  $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$  be a linear map. Then  $\exists! A \in \mathbb{F}^{m \times n}$  s.t.  $f(\vec{x}) = A\vec{x} \quad \forall \vec{x} \in \mathbb{F}^n$ . Conversely, if  $A \in \mathbb{F}^{m \times n}$  then  $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$  is a linear map.

## Gaussian Elimination

To solve the system  $A\vec{x} = \vec{b}$ , we can use Gaussian Elimination and backward substitution.

$$A\vec{x} = \vec{b} \rightarrow \underbrace{(A|\vec{b})}_{\text{aug. matrix}} \xrightarrow{\text{E.R.O.}} \underbrace{(U|\vec{c})}_{\text{upper } \Delta} \rightarrow \vec{x} = U^{-1}\vec{c} \xrightarrow{\text{backward substitution}}$$

We use elementary row operations to do Gaussian Elimination

- Add a multiple of a row to another
- Exchange two rows
- Multiply a row by a non-zero scalar

## Elementary Matrices

- $E_{ij}(\alpha)$ : add  $\alpha$  times row  $j$  to row  $i$ ,  $i \neq j$
- $E_{ij}$ : exchange row  $i$  and  $j$ ,  $i \neq j$
- $E_i(\alpha)$ : multiply row  $i$  by  $\alpha \neq 0$

Each elementary matrix is invertible and the inverse is of the same type

- $E_{ij}(\alpha)^{-1} = E_{ij}(-\alpha)$
- $E_{ij}^{-1} = E_{ij}$
- $E_i(\alpha)^{-1} = E_i\left(\frac{1}{\alpha}\right)$

## LU Factorization

If  $A$  is an  $n \times n$  matrix, invertible and can be transformed into an upper-triangular matrix  $U$  using only elementary row operations of type I

$$E_k \cdots E_2 E_1 A = U \Rightarrow A = LU, L = (E_k \cdots E_2)^{-1} = E_1^{-1} \cdots E_k^{-1}$$

To solve a linear system using LU factorization, we have

$$A\vec{x} = \vec{b} \Leftrightarrow LU\vec{x} = \vec{b} = \begin{cases} L\vec{z} = \vec{b} & \text{forward sub.} \\ U\vec{z} = \vec{b} & \text{back. sub.} \end{cases}$$

$$\text{We also get } \det(A) = \prod_{i=1}^n u_{ii}$$

Proposition: Let  $A$  be invertible. If  $A = LU$  then this decomposition is unique.

## Pivoting and Permutation

Definition: A matrix obtained from  $I$  by any row interchanges is called a permutation matrix

Equivalent Definition:  $A$  is a permutation matrix if each row and each column has exactly one 1, all the other entries being 0.

Proposition: Let  $P$  be a permutation matrix, then  $P$  is invertible and  $P^{-1} = P^T$ .

Theorem: Any invertible matrix  $A$  has a permuted LU factorization, namely  $PA = LU$

$P$ : permutation matrix

$L$ : lower unitriangular matrix

$U$ : Upper triangular matrix

We use the permuted LU factorization because sometimes row interchanges are necessary if the pivot is 0

## Partial Pivoting

Even if not needed, it is a good idea to use row interchanges

Let  $A^{(k-1)}$  be the matrix at step  $k$  of Gaussian elimination

Pick  $j$  s.t.  $|a_{jk}^{(k-1)}| = \max_{k \leq i \leq n} |a_{ik}^{(k-1)}|$  and do  
 $R_j \leftrightarrow R_k$

## LDV Factorization

Let  $A$  be invertible and assume  $A = LU$

$$A = LU = LDV \quad D = \text{diag}(u_{11}, u_{22}, \dots, u_{nn}) \quad V = D^{-1}U \text{ (upper uni-0)}$$

Remark: i) If  $A = LDV$ , then  $A^T = (LDV)^T = V^T D^T L^T$   
ii) If  $A^T = A$ , then  $A = LDL^T = V^T D V^T$

If  $A = LDV$  then  
 $A\vec{x} = \vec{b} \Leftrightarrow LDV\vec{x} = \vec{b} \Leftrightarrow \begin{cases} L\vec{z} = \vec{b} & \text{forward substitution} \\ D\vec{z} = \vec{y} & \text{"rescaling" } z_i = \frac{1}{d_{ii}} y_i \\ V\vec{y} = \vec{b} & \text{backward substitution} \end{cases}$

For permuted LDV factorization, we have  $PA = LDV$

## General Linear Systems

We can generalize LU decomposition to general matrix  $A$

$L$  is square lower unitriangular

$U$  is in row echelon form where:

- all nonzero rows are above zero rows
- for each nonzero row, the leading entry is strictly on the right of the leading entries of rows above it

Example:  $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$  is in row echelon form

# Inner Product

Let  $V$  be a vector space over  $\mathbb{F}$ . A map  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{F}$  is said to be an inner product if

$$i) \langle \alpha\vec{u} + \beta\vec{v}, \vec{w} \rangle = \alpha\langle \vec{u}, \vec{w} \rangle + \beta\langle \vec{v}, \vec{w} \rangle$$

$$\forall \vec{u}, \vec{v}, \vec{w} \in V \quad \forall \alpha, \beta \in \mathbb{F}$$

$$ii) \langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle} \quad \forall \vec{u}, \vec{v} \in V$$

$$iii) \langle \vec{u}, \vec{u} \rangle \geq 0 \quad \forall \vec{u} \in V$$

$$\langle \vec{u}, \vec{u} \rangle = 0 \text{ iff } \vec{u} = \vec{0}$$

From i) and ii), we get

$$\langle \vec{u}, \alpha\vec{u} + \beta\vec{v} \rangle = \bar{\alpha}\langle \vec{u}, \vec{u} \rangle + \bar{\beta}\langle \vec{u}, \vec{v} \rangle$$

## Vector Norm

Let  $V$  be a vector space over  $\mathbb{F}$ . A map  $\|\cdot\|: V \rightarrow \mathbb{R}$  is said to be a norm if

$$i) \|\vec{u}\| \geq 0 \quad \forall \vec{u} \in V \text{ and } \|\vec{u}\|=0 \text{ iff } \vec{u} = \vec{0}$$

$$ii) \|\alpha\vec{u}\| = |\alpha| \|\vec{u}\| \quad \forall \vec{u} \in V \quad \forall \alpha \in \mathbb{F}$$

$$iii) \|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| \quad \forall \vec{u}, \vec{v} \in V$$

Any inner product induces a norm

$$\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}, \quad \vec{u} \in V$$

Cauchy-Schwarz inequality: Let

$\langle \cdot, \cdot \rangle$  be an inner product on  $V$  and let  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ . Then,

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\| \quad \forall \vec{u}, \vec{v} \in V$$

P-norm: Let  $1 \leq p \leq \infty$ . The P-norm on  $\mathbb{F}^n$  is defined by

$$\text{if } p=\infty: \|\vec{x}\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad \vec{x} \in \mathbb{F}^n$$

$$\text{if } p=\infty: \|\vec{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

## Matrix Norm

Frobenius norm: Let  $A=(a_{ij}) \in \mathbb{F}^{m,n}$

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

Induced Matrix Norm: Let  $\|\cdot\|_1$  and  $\|\cdot\|_2$  be norms on  $\mathbb{F}^m$  and  $\mathbb{F}^n$ , respectively. Then

$$\|A\|_{1,2} = \sup_{\substack{\vec{x} \in \mathbb{F}^n \\ \vec{x} \neq \vec{0}}} \frac{\|A\vec{x}\|_2}{\|\vec{x}\|_1} = \sup_{\substack{\vec{x} \in \mathbb{F}^n \\ \vec{x} \neq \vec{0}}} \|A\vec{x}\|_2$$

Proposition: For any induced matrix norm, we have

$$i) \|A\vec{x}\| \leq \|A\| \|\vec{x}\| \quad \forall \vec{x} \in \mathbb{F}^n \quad \forall A \in \mathbb{F}^{m,n}$$

$$ii) \|I\| = 1$$

$$iii) \|AB\| \leq \|A\| \|B\| \quad \forall A \in \mathbb{F}^{m,k} \quad \forall B \in \mathbb{F}^{k,n}$$

P-norm: let  $A \in \mathbb{F}^{m,n}$  then

$$\forall 1 \leq p \leq \infty, \quad \|A\|_p = \sup_{\substack{\vec{x} \in \mathbb{F}^n \\ \vec{x} \neq \vec{0}}} \frac{\|A\vec{x}\|}{\|\vec{x}\|}$$

Special cases for  $p=1$  and  $p=\infty$

$$\|A\|_1 = \max_{1 \leq j \leq n} \left\{ \sum_{i=1}^m |a_{ij}| \right\}$$

$$\|A\|_\infty = \max_{1 \leq i \leq m} \left\{ \sum_{j=1}^n |a_{ij}| \right\}$$

## Conditioning

Let  $\|\cdot\|$  denote the vector norm and its induced matrix norm. Then for any nonsingular matrix  $A$  we define

$$K(A) = \|A\| \|A^{-1}\|$$

the condition number of  $A$

Properties:

- i)  $K(A) \geq 1$
- ii)  $A \in \mathbb{F}^n, \alpha \neq 0 \Rightarrow K(\alpha A) = K(A)$

If  $K(A) \gg 1$ , we say  $A$  is ill-conditioned (well-conditioned if  $K(A)$  "close" to 1)

$K(A)$  depends on norm:  $K_1(A), K_2(A), K_\infty(A)$

## System Perturbation

Exact system:  $A\vec{x} = \vec{b}$  (1)

Perturbed system:  $A\vec{x} = \vec{b} + \vec{e}$  (2)

Theorem: Let  $\vec{x}$  be the solution to (1) with  $\vec{b} \neq \vec{0}$  and let  $\vec{x}'$  be the solution to (2). Then

$$\frac{1}{K(A)} \frac{\|\vec{e}\|}{\|\vec{b}\|} \leq \frac{\|\vec{x}' - \vec{x}\|}{\|\vec{x}\|} \leq K(A) \frac{\|\vec{e}\|}{\|\vec{b}\|}$$

## Positive Definite Matrices

A symmetric matrix  $K \in \mathbb{F}^{n,n}$  is said to be positive definite if

$$\vec{x}^\top K \vec{x} > 0 \quad \forall \vec{x} \in \mathbb{F}^n, \vec{x} \neq \vec{0}$$

Theorem: Any inner product on  $\mathbb{F}^n$  is of the form

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^\top K \vec{y}$$

for some symmetric positive definite matrix  $K$

Remark:  $K$  positive definite  $\Rightarrow \ker(K) = \{\vec{0}\}$

## Quadratic Forms

A function  $q: \mathbb{F}^n \rightarrow \mathbb{F}$  of the form

$$q(\vec{x}) = q(x_1, x_2, \dots, x_n) = \sum_{i,j} c_{ij} x_i x_j$$

is called a quadratic form.

Proposition: For any quadratic form on  $\mathbb{F}^n$ , there exists a unique symmetric matrix  $K \in \mathbb{F}^{n,n}$  s.t.  $q(\vec{x}) = \vec{x}^\top K \vec{x}$

## Determine Positive Definite

Method 1: Complete the square

Let  $K \in \mathbb{F}^{n,n}$  be symmetric

Then take its quadratic form  $q(\vec{x}) = \vec{x}^\top K \vec{x}$  and expand it. Then complete the square.

If the quadratic form is strictly greater than 0 for all  $\vec{x} \neq \vec{0}$  and 0 for  $\vec{x} = \vec{0}$ , then  $K$  is positive definite.

## Continued:

For a general  $2 \times 2$  matrix, let

$$K = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Then  $K$  is positive definite iff  $a > 0$  and  $c - \frac{b^2}{a} > 0$

## Method 2: LDL<sup>T</sup> factorization

Theorem: A matrix  $A \in \mathbb{F}^{n,n}$  is symmetric positive definite if and only if  $A = LDL^T$  with  $L$  lower unit  $\Delta$  and  $D = \text{diag}(d_1, \dots, d_n)$  with  $d_i > 0 \forall i \in \mathbb{N}$

## Conjugate Transpose

Definition: Let  $A \in \mathbb{C}^{n,n}$ . The conjugate transpose of  $A$ , written  $A^H$ , is

$$A^H = (\bar{a}_{ij}) \in \mathbb{C}^{n,n}$$

$$A = (a_{ij}), \quad A^H = (\bar{a}_{ji})$$

Properties:

- i)  $(A+B)^H = A^H + B^H$
- ii)  $(\alpha A)^H = \bar{\alpha} A^H$
- iii)  $(A^H)^H = A$
- iv)  $(AB)^H = B^H A^H$

Definition: A matrix  $A \in \mathbb{C}^{n,n}$  is said to be Hermitian if  $A^H = A$

If  $A = A^H$

$$\cdot a_{ii} \in \mathbb{R}$$

$$\cdot x^H A x \in \mathbb{R} \quad \forall x \in \mathbb{C}^n$$

## Positive Definite Matrices: Complex Case

Definition: A Hermitian matrix  $A \in \mathbb{C}^{n,n}$  is said to be positive definite if

$$x^H A x > 0 \quad \forall x \in \mathbb{C}^n, x \neq \vec{0}$$

For any inner product on  $\mathbb{C}^n$   $\langle \cdot, \cdot \rangle$  can be written  $\langle x, y \rangle = x^H K y$  for some  $K$  SPD

•  $A$  is HPD if  $A = LDL^H$ ,  $L$  lower unitriangular,  $D = \text{diag}(d_1, \dots, d_n) \in \mathbb{C}^{n,n}$ ,  $d_i > 0$

## Cholesky Factorization

If  $A$  is HPD, then we can get the following decomposition:

$$A = LDL^H = (L \sqrt{D}) (\sqrt{D} L^H) = M M^H$$

where  $M = L \sqrt{D}$ ,  $\sqrt{D} = \text{diag}(\sqrt{d_1}, \dots, \sqrt{d_n})$

This is called the Cholesky factorization

In practice, we don't go through Gaussian Elimination to get  $M M^H$

We can find  $m_{ij}$  using this algorithm for  $j=1, \dots, n$  do

$$m_{jj} = \sqrt{a_{jj} - \sum_{i=1}^{j-1} m_{ij} m_{ij}}$$

for  $i=j+1, \dots, n$  do

$$m_{ij} = \frac{a_{ij}}{m_{jj}} \left( a_{jj} - \sum_{k=1}^{j-1} m_{kj} m_{kj} \right)$$

End for

end for

## Orthogonal and Orthonormal Basis

$V$  vector space over  $\mathbb{F}$ ,  $\langle \cdot, \cdot \rangle$  inner product,  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$

Definition: we say that  $\vec{u}, \vec{v} \in V$  are orthogonal if  $\langle \vec{u}, \vec{v} \rangle = 0$

we say that a set  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is

• orthogonal if  $\langle v_i, v_j \rangle = 0 \quad \forall i \neq j$

• orthonormal if  $\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

Proposition: If  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is an orthogonal set of nonzero vectors, then they are linearly independent

Proposition: Let  $\{\vec{v}_1, \dots, \vec{v}_n\}$  be an orthogonal basis of a subspace  $W$  of  $V$ .

Then  $\forall \vec{w} \in W$ , we have

$$\vec{w} = \sum_{i=1}^n \frac{\langle \vec{w}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$$

If it is an ONB, then  $\vec{w} = \sum_{i=1}^n \frac{\langle \vec{w}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$

## Gram-Schmidt

The Gram-Schmidt algorithm transforms any basis  $\{\vec{w}_1, \dots, \vec{w}_n\}$  into an orthogonal basis

$$\begin{aligned} \vec{v}_1 &= \vec{w}_1 \\ \vec{v}_2 &= \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 \\ \vec{v}_3 &= \vec{w}_3 - \frac{\langle \vec{w}_3, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 - \frac{\langle \vec{w}_3, \vec{v}_2 \rangle}{\|\vec{v}_2\|^2} \vec{v}_2 \\ &\vdots \\ \vec{v}_n &= \vec{w}_n - \sum_{i=1}^{n-1} \frac{\langle \vec{w}_n, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i \end{aligned}$$

We can turn  $\{\vec{v}_1, \dots, \vec{v}_n\}$  into an ONB by setting  $\vec{u}_i = \frac{1}{\|\vec{v}_i\|} \vec{v}_i$

## QR Decomposition

Let  $V = \mathbb{F}^n$ .  $\{\vec{w}_1, \dots, \vec{w}_n\}$  a basis and  $\{\vec{v}_1, \dots, \vec{v}_n\}$  an orthogonal basis

Then we get

$$\underbrace{(\vec{w}_1 | \vec{w}_2 | \dots | \vec{w}_n)}_{A \in \mathbb{F}^{n,n}} = (\vec{v}_1 | \vec{v}_2 | \dots | \vec{v}_n) \underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}}_{\text{orthogonal columns}}$$

The QR factorization is

$$A = QR$$

where  $A = (\vec{w}_1 | \vec{w}_2 | \dots | \vec{w}_n)$

$$Q = (\vec{v}_1 | \vec{v}_2 | \dots | \vec{v}_n)$$

$$R = \begin{pmatrix} \|\vec{v}_1\| & & & \\ 0 & \|\vec{v}_2\| & & \\ & 0 & \ddots & \\ & & 0 & \|\vec{v}_n\| \end{pmatrix}$$

Some remarks:

For  $A \in \mathbb{F}^{n,n}$  with  $n$  columns we have  $A = \underbrace{QR}_{m \times n}$

$$i) Q^H Q = I_n$$

ii)  $R$  is upper triangular and invertible

$$iii) A^H A = R^H R \quad (\text{Cholesky for } A^H A)$$

## Unitary (Orthogonal) Matrices

Definition: A matrix  $Q \in \mathbb{C}^{n,n}$  is called unitary if  $Q^H Q = Q Q^H = I_n$ .

Proposition: Let  $Q$  and  $S$  be unitary (orthogonal) matrices of size  $n \times n$ . Then

$$i) \text{The columns of } Q \text{ is an ONB for } \mathbb{C}^n$$

$$ii) |\det(Q)| = 1$$

$$iii) QS \text{ is unitary (orthogonal)}$$

$$iv) \langle Q\vec{z}, Q\vec{y} \rangle = \langle \vec{z}, \vec{y} \rangle \quad \forall \vec{z}, \vec{y} \in \mathbb{C}^n$$

## Orthogonal Complement

Let  $\{\vec{v}_1, \dots, \vec{v}_n\}$  be an orthogonal basis for  $W$  subspace of  $V$ . The orthogonal complement of  $W$  in  $V$  is

$$W^\perp = \{\vec{v} \in V, \langle \vec{v}, \vec{w} \rangle = 0 \forall \vec{w} \in W\}$$

## Orthogonal Projection

Let  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$  be an orthogonal basis for  $W$  subspace of  $V$ . The orthogonal projection of  $\vec{v} \in V$  onto  $W$  is

$$\text{Proj}_W(\vec{v}) = \sum_{i=1}^n \frac{\langle \vec{v}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$$

If  $W = \{0\}$  then we set  $\text{Proj}_W(\vec{v}) = \vec{0}$

Remarks:

- $\vec{v} \in W \iff \vec{v} = \text{Proj}_W(\vec{v})$
- $\vec{v}_k = \vec{w}_k - \sum_{i=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i = \vec{w}_k - \text{Proj}_{\text{span}\{\vec{v}_1, \dots, \vec{v}_k\}}(\vec{w}_k)$

Proposition: Let  $W$  be a subspace of  $V$  and let  $\vec{p} = \text{Proj}_W(\vec{v})$  for some  $\vec{v}$ . Then

- i)  $\vec{p} \in W$  and  $\vec{v} - \vec{p} \in W^\perp$
- ii)  $\|\vec{v} - \vec{p}\| < \|\vec{v} - \vec{w}\| \forall \vec{w} \in W, \vec{w} \neq \vec{p}$

## Projection Matrix

$V = \mathbb{F}^m$ , let  $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$  be an orthonormal basis of  $W \subseteq V$  and  $Q = (\vec{u}_1 | \dots | \vec{u}_n)$

Then for any  $\vec{v} \in V$ , we get

$$QQ^H \vec{v} = \text{Proj}_W(\vec{v})$$

The matrix  $P = QQ^H$  is called a projection matrix and satisfies

- i)  $P$  is Hermitian
- ii)  $P^2 = P$  ( $P$  idempotent)

## Solving Linear System using QR decom.

Given a system  $A\vec{x} = \vec{b}$

Then

$$\begin{aligned} A\vec{x} = \vec{b} &\iff QR\vec{x} = \vec{b} \\ &\Rightarrow Q^H QR\vec{x} = Q^H \vec{b} \\ &\Rightarrow R\vec{x} = Q^H \vec{b} \end{aligned}$$

$\vec{x} = R^{-1}Q^H \vec{b}$  is not always solution to the system  $A\vec{x} = \vec{b}$

- If  $Q^H \vec{b} = \vec{b}$ , namely  $\vec{b} \in \text{col}(A)$ , then  $\vec{x} = R^{-1}Q^H \vec{b}$  is solution to  $A\vec{x} = \vec{b}$
- If  $Q^H \vec{b} \neq \vec{b}$ , namely  $\vec{b} \notin \text{col}(A)$ , then  $A\vec{x} = \vec{b}$  has no solution

## Least Squares Minimization

Let  $A \in \mathbb{C}^{m,n}$  with rank  $n$ . The least squares minimization problem is when the system  $A\vec{x} = \vec{b}$  has no solutions and we want to find  $\vec{x}$  that solves

$$\min \|A\vec{x} - \vec{b}\|_2^2$$

We want to find  $\vec{y} = A\vec{x} \in \text{col}(A)$  that is closest to  $\vec{b}$ .

We have  $\vec{y} = \text{Proj}_{\text{col}(A)} \vec{b}$ . Therefore, we can find the solution to the least squares minimization problem using the reduced QR factorization.

$$A\vec{x} = Q\vec{z} \iff R\vec{z} = Q^H \vec{b}$$

$\vec{z}$  is the least squares solution to  $A\vec{x} = \vec{b}$

## Eigenvalues and Eigenvectors

Let  $A \in \mathbb{C}^{n,n}$ . Recall that a scalar  $\lambda$  is an eigenvalue if  $A\vec{v} = \lambda\vec{v}$  for some nonzero vector  $\vec{v} \neq \vec{0}$ . Then  $\vec{v}$  is called an eigenvector corresponding to  $\lambda$ . We call  $(\lambda, \vec{v})$  an eigenpair.

To compute  $(\lambda, \vec{v})$ , we have

$$A\vec{v} = \lambda\vec{v} \iff (A - \lambda I)\vec{v} = 0$$

has a nonzero solution iff  $A - \lambda I$  is singular, namely  $\det(A - \lambda I) = 0$

We define the characteristic polynomial

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

$$= (\lambda_1 - \lambda)^{m_1} (\lambda_2 - \lambda)^{m_2} \cdots (\lambda_n - \lambda)^{m_n}$$

where  $\lambda_1, \lambda_2, \dots, \lambda_n$  are distinct eigenvalues and  $m_{\lambda_j}$  is the algebraic multiplicity of  $\lambda_j$

$V_{\lambda_j}$  is the eigenspace corresponding to  $\lambda_j$ , defined by  $V_{\lambda_j} = \ker(A - \lambda_j I)$ , where  $\dim(V_{\lambda_j})$  is the geometric multiplicity of  $\lambda_j$

We have  $m_{\lambda_1} + m_{\lambda_2} + \dots + m_{\lambda_n} = n$  and  $1 \leq \dim(V_{\lambda_j}) \leq m_{\lambda_j}$

Proposition: Let  $A \in \mathbb{C}^{n,n}$

- i)  $A$  is singular  $\iff \lambda = 0$  is an eigenvalue
- ii)  $A$  and  $A^T$  have the same eigenvalues, but different eigenvectors in general
- iii) If  $A$  is real and  $(\lambda, \vec{v}) = (\alpha i\beta, \vec{u} + i\vec{v})$  is an eigenpair, then  $(-\alpha i\beta, \vec{u} - i\vec{v})$  is also an eigenpair.
- iv)  $\det(A) = \prod_{i=1}^n \lambda_i$  and  $\text{tr}(A) = \sum_{i=1}^n \lambda_i$ , where  $\lambda_i$  are eigenvalues of  $A$ .
- v) For every induced matrix norm,  $\|A\|_i \leq \|A\|_1$  for all  $i = 1, \dots, n$ . In particular, we have

$$p(A) = \max_{1 \leq j \leq n} |\lambda_j| \leq \|A\|_1$$

where  $p(A)$  is the spectral radius of  $A$ .

## Invariant Subspace

We say that a subspace  $W$  of a vector space  $V$  is invariant from a linear map  $L: V \rightarrow V$  if  $L(v) \in W \forall v \in W$

For a matrix  $A \in \mathbb{C}^{n,n}$ , we say that  $W \subseteq V$  is invariant for  $A$  if  $Aw \in W \forall w \in W$

$$Aw \in W \quad \forall w \in W$$

## Gershgorin Circle Theorem

Definition: For  $i=1, 2, \dots, n$ , the  $i$ th Gershgorin disk is defined by

$$D_i = D(a_{ii}, r_i) = \{z \in \mathbb{C}, |z - a_{ii}| \leq r_i\},$$

$$r_i = \sum_{j \neq i} |a_{ij}|$$

We call  $D_A = \bigcup_{i=1}^n D_i$  the Gershgorin domain.

Theorem: All the eigenvalues of  $A$  lies in its Gershgorin domain  $D_A$ .

Theorem: Let  $S_1$  be the union of  $k$  Gershgorin disks of  $A$  and let  $S_2 = D_A \setminus S_1$ , the union of the  $n-k$  remaining ones. If  $S_1, S_2 = \emptyset$ , then  $S_1$  contains  $k$  eigenvalues and  $S_2$  contains  $n-k$  eigenvalues.

## Strictly Diagonally Dominant Matrices

A square matrix  $A$  is said to be strictly diagonally dominant (by row) if  $|a_{ii}| > \sum_{j \neq i} |a_{ij}| = r_i \quad \forall i = 1, 2, \dots, n$

Theorem: If  $A$  is strictly diagonally dominant, then  $A$  is invertible.

## Similar Matrices

Definition: Two matrices  $A$  and  $B$  are said to be similar if there exists an invertible matrix  $S$  st.

$$S^{-1}AS = B \quad (A = SBS^{-1})$$

Two similar matrices have the same eigenvalues

## Diagonalization

A square matrix is said to be diagonalizable if it is similar to a diagonal matrix.

$$S^{-1}AS = \Lambda$$

$\Lambda$  is a diagonal matrix with the eigenvalues of  $A$  on the diagonal.

$S$  has the eigenvectors corresponding to the eigenvalues on its columns.

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

$$S = (\vec{v}_1 | \vec{v}_2 | \cdots | \vec{v}_n)$$

$(\lambda_i, \vec{v}_i)$  eigenpairs

$S^{-1}$  has the left eigenvectors of  $A$  corresponding to the eigenvalues on its rows.

Theorem: Let  $A \in \mathbb{C}^{n,n}$ . The following statements are equivalent:

- i) There exists a basis of  $\mathbb{C}^n$  consisting of eigenvectors of  $A$
- ii)  $m_{\lambda_i} = \dim(V_{\lambda_i})$  for all eigenvalues  $\lambda_i$  of  $A$
- iii)  $A$  is diagonalizable

Corollary: If all the eigenvalues of  $A$  are distinct, then  $A$  is diagonalizable.

Remark: If  $A \in \mathbb{R}^{n,n}$  and all the eigenvalues are real and satisfy  $m_{\lambda} = \dim(V_{\lambda})$ , then we can take  $S, A \in \mathbb{R}^{n,n}$  and we say  $A$  is diagonalizable over  $\mathbb{R}$ .

## Exponentials of Matrices

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \dots$$

If  $A$  is diagonalizable, we have

$$A = S \Lambda S^{-1}$$

$$e^A = S \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) S^{-1}$$

## Eigenvalues and Eigenvectors of Hermitian matrices

Proposition: Let  $A \in \mathbb{C}^{n,n}$ . If  $A$  is Hermitian then all the eigenvalues are real.

Proposition: Let  $A \in \mathbb{C}^{n,n}$ . If  $A$  is Hermitian then eigenvectors corresponding to distinct eigenvalues are orthogonal.

## Unitarily Diagonalizable Matrices

A matrix  $A \in \mathbb{C}^{n,n}$  is said to be unitarily diagonalizable if  $\exists U \in \mathbb{C}^{n,n}$  unitary and  $A = U \Lambda U^{-1}$  diagonal s.t.  $U^H A U = \Lambda$ .

Spectral Theorem: If  $A$  is Hermitian then  $A$  is unitarily diagonalizable.

Principal Axes Theorem: Let  $A \in \mathbb{C}^{n,n}$ . The following are equivalent

- (a)  $A$  has an orthonormal set of  $n$  eigenvectors in  $\mathbb{R}^n$ .
- (b)  $A$  is orthogonally diagonalizable.
- (c)  $A$  is symmetric.

## Positive Definite Matrices (eigenvalues)

Theorem: A Hermitian Matrix is Positive definite  $\iff$  all eigenvalues are positive.

## 2-norm of Hermitian Matrices

Proposition: Let  $A \in \mathbb{C}^{n,n}$ . For any unitary matrix  $U \in \mathbb{C}^{n,n}$  we have  $\|AU\|_2 = \|A\|_2 = \|AU\|_F$

Proposition: Let  $A \in \mathbb{C}^{n,n}$  be Hermitian and let  $\lambda_1, \dots, \lambda_n$  be its eigenvalues. Then

$$\|A\|_2 = \max_{1 \leq i \leq n} |\lambda_i| = p(A)$$

We also get that if a matrix  $A$  is Hermitian positive definite, we get

$$\|A\|_2 = \lambda_{\max}, \quad \|A^{-1}\|_2 = \frac{1}{\lambda_{\min}}, \quad k(A) = \frac{\lambda_{\max}}{\lambda_{\min}}$$

## Raleigh Quotient

Let  $A \in \mathbb{C}^{n,n}$  be Hermitian

Let  $r_A: \mathbb{C}^n \setminus \{\vec{0}\} \rightarrow \mathbb{R}$  be defined by

$$r_A(\vec{v}) = \frac{\langle A\vec{v}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}$$

We can use this to find an eigenvalue  $\lambda_i$  corresponding to a given eigenvector  $\vec{v}_i$  of  $A$ .

$$\lambda_i = \frac{\vec{v}_i^H A \vec{v}_i}{\vec{v}_i^H \vec{v}_i} = \frac{\langle A\vec{v}_i, \vec{v}_i \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle} = r_A(\vec{v}_i)$$

We also have  $\lambda_1 \leq r_A(\vec{v}) \leq \lambda_n \quad \forall \vec{v} \in \mathbb{C}^n \setminus \{\vec{0}\}$  and  $\exists \vec{v}_i, \vec{v}_n$  s.t.  $r_A(\vec{v}_i) = \lambda_i$  and  $r_A(\vec{v}_n) = \lambda_n$ .

## Schur Decomposition

Theorem: Let  $A \in \mathbb{C}^{n \times n}$ . Then there exists a unitary matrix such that

$$U^H A U = T \quad (A = U T U^H)$$

where  $T \in \mathbb{C}^{n \times n}$  is upper triangular.  
Moreover, the diagonals of  $T$  are eigenvalues of  $A$ .

## Normal Matrices

A matrix  $A$  is said to be normal if  $A^H A = A A^H$

Theorem: Let  $A \in \mathbb{C}^{n \times n}$ . Then  $A$  is normal if and only if it is unitarily diagonalizable.

## Jordan Matrix

Given a scalar  $\lambda$  and an integer  $n$ , the Jordan block  $J_{\lambda,n}$  is the  $n \times n$  matrix with  $\lambda$  on the diagonal, 1 on the upper diagonal, and 0 everywhere else, namely

$$J_{\lambda,1} = (\lambda), \quad J_{\lambda,2} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad J_{\lambda,3} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

and so on

A Jordan matrix is a block diagonal matrix of the form

$$J = \text{diag}(J_{\lambda_1, n_1}, J_{\lambda_2, n_2}, \dots, J_{\lambda_k, n_k})$$

$$= \begin{pmatrix} J_{\lambda_1, n_1} & & & \\ & J_{\lambda_2, n_2} & & \\ & & \ddots & \\ 0 & & & J_{\lambda_k, n_k} \end{pmatrix}$$