MAT 3341

Applied Linear Algebra

Study Guide

winter 2024

$Ax = b$ $\qquad$ $Ax = \lambda x$

$A = LU$ $\qquad$ $PA = LU$ $\qquad$ $A = QR$

Matrix Arithmetic

- addition / subtraction:
  $A=(a_{ij})$ $B=(b_{ij})$ $m\times n$
  $A\pm B = (a_{ij}\pm b_{ij})$
- scalar multiplication:
  $A=(a_{ij})$ α is a scalar
  $\alpha A = (\alpha a_{ij})$
- matrix-vector multiplication:
  $A=(a_1\ a_2 \cdots a_n)$
  $a_1, a_2, \ldots, a_n \in \mathbb{F}^m$
  $x \in \mathbb{F}^n$
  $Ax = \sum_{i=1}^{n} x_i a_i \in \mathbb{F}^m$
- matrix-matrix multiplication
  A : $m\times q$ matrix
  B : $q\times n$ matrix with columns $b_1, b_2, \ldots, b_n \in \mathbb{F}^q$
  $AB = (Ab_1\ Ab_2 \cdots Ab_n)$
  or if $C=AB$, we have that
  $c_{ij} = \sum_{k=1}^{q} a_{ik}b_{kj}$

2 special Matrices

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

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

For a $m\times n$ matrix A:
- $A+0 = 0+A = A$
- $I_m A = A I_n = A$

Diagonal Matrix

$A = \operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & & \alpha_n \end{pmatrix} \in \mathbb{F}^{n\times n}$

Transpose

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

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

$A^T = A \Rightarrow A$ 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 $f: \mathbb{F}^n \to \mathbb{F}^m$ is a linear map if:
- $f(\vec{x}+\vec{y}) = f(\vec{x}) + f(\vec{y})$ $\forall \vec{x}, \vec{y} \in \mathbb{F}^n$
- $f(\alpha\vec{x}) = \alpha f(\vec{x})$ $\forall \alpha \in \mathbb{F}, \vec{x} \in \mathbb{F}^n$

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

Gaussian Elimination

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

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

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 α 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

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

To solve a linear system using LU factorization, we have

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

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 $A = LU$

$A = LU = LDV$ $D = \operatorname{diag}(u_{11}, u_{22}, \ldots, u_{nn})$ $V = D^{-1}U$ (upper uni-Δ)

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

If $A = LDV$ then

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

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 $\langle\cdot,\cdot\rangle: V\times V \to \mathbb{F}$ is said to be an inner product if

i) $\langle \alpha\vec{u}+\beta\vec{v}, \vec{w}\rangle = \alpha\langle\vec{u},\vec{w}\rangle + \beta\langle\vec{v},\vec{w}\rangle$ $\quad \forall \vec{u},\vec{v},\vec{w}\in V \;\; \forall \alpha,\beta\in\mathbb{F}$

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

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

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

From i) and ii), we get

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

Vector Norm

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

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

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

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

Any inner product induce a norm

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

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

- if $p<\infty$: $\|\vec{x}\|_p = \left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}, \; \vec{x}\in\mathbb{F}^n$
- if $p=\infty$: $\|\vec{x}\|_\infty = \max_{1\leq i\leq n}|x_i|$

Matrix Norm

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

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

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

$$\|A\|_{a,b} = \sup_{\substack{\vec{x}\in\mathbb{F}^n\\ \vec{x}\neq\vec{0}}} \frac{\|A\vec{x}\|_b}{\|\vec{x}\|_a} = \sup_{\|\vec{x}\|_a=1}\|A\vec{x}\|_b$$

Proposition: For any induced matrix norm, we have

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

ii) $\|I_n\|=1$

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

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

$\forall 1\leq p\leq\infty, \; \|A\|_p = \sup_{\vec{x}\neq\vec{0}}\frac{\|A\vec{x}\|}{\|\vec{x}\|}$

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:

- $K(A)\geq 1$
- $\forall\alpha\in\mathbb{F}, \alpha\neq 0 \;\; K(\alpha A) = K(A)$

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: $A\vec{x}=\vec{b}$ $\quad (*)$

Perturbed system: $A\tilde{\vec{x}} = \vec{b}+\vec{\delta b}$ $\quad (**)$

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

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

Positive Definite Matrices

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

$$\vec{x}^TK\vec{x}>0 \quad \forall\vec{x}\in\mathbb{R}^n, \; \vec{x}\neq\vec{0}$$

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

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

for some symmetric positive definite matrix K

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

Quadratic Forms

A function $q:\mathbb{R}^n\to\mathbb{R}$ of the form

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

is called a quadratic form.

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

Determine Positive Definite

Method 1: Complete the square

Let $K\in\mathbb{R}^{n\times n}$ be symmetric

Then take its quadratic form $q(\vec{x})=\vec{x}^TK\vec{x}$ 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 2×2 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^T factorization

Theorem: A matrix $A\in\mathbb{R}^{n\times n}$ is symmetric positive definite if and only if $A=LDL^T$ with L lower uni Δ and $D=\mathrm{diag}(d_{11},\dots,d_{nn})$ with $d_{ii}>0 \;\; \forall 1\leq i\leq n$

Conjugate Transpose

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

$$A^H = (\bar{A})^T\in\mathbb{C}^{n\times m}$$

$A=(a_{ij})$, $A^H=(\overline{a_{ji}})$

Properties:

- $(A+B)^H = A^H+B^H$
- $(\alpha A)^H = \bar{\alpha}A^H$
- $(A^H)^H = A$
- $(AB)^H = B^HA^H$

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

If $A=A^H$

- $a_{ii}\in\mathbb{R}$
- $\vec{x}^HA\vec{x}\in\mathbb{R} \;\; \forall\vec{x}\in\mathbb{C}^n$

Positive Definite Matrices: Complex Case

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

$$x^HAx>0 \quad \forall x\in\mathbb{C}^n, \; x\neq 0$$

- For any inner product on $\mathbb{C}^n$ $\langle\cdot,\cdot\rangle$ can be written $\langle x,y\rangle = x^HKy$ for some K SPD
- A is HPD if $A=LDL^H$, L lower unitriangular, $D=\mathrm{diag}(d_{11},\dots,d_{nn})\in\mathbb{R}^{n\times n}$, $d_{ii}>0$

Cholesky Factorization

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

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

where $M=L\sqrt{D}$, $\sqrt{D}=\mathrm{diag}(\sqrt{d_{11}},\dots,\sqrt{d_{nn}})$

This is called the Cholesky factorization

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

We can find m_{ij} using this algorithm

```
for j=1,...,n do
    m_jj = sqrt(a_jj - sum_{k=1}^{j-1} m_jk conj(m_jk))
    for i=j+1,...,n do
        m_ij = (1/m_jj) (a_ij - sum_{k=1}^{j-1} m_ik conj(m_jk))
    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 $\{v_1,\dots,v_n\}$ is

- orthogonal if $\langle v_i,v_j\rangle=0 \;\; \forall i\neq j$
- 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 $\vec{w} = \sum_{i=1}^{n}\langle\vec{w},\vec{v}_i\rangle\vec{v}_i$

Gram-Schmidt

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

Goal: Transform a basis $\{\vec{w}_1,\dots,\vec{w}_n\}$ into an orthogonal basis $\{\vec{v}_1,\dots,\vec{v}_n\}$

- $\vec{v}_1 = \vec{w}_1$
- $\vec{v}_2 = \vec{w}_2 - \frac{\langle\vec{w}_2,\vec{v}_1\rangle}{\|\vec{v}_1\|^2}\vec{v}_1$
- $\vec{v}_3 = \vec{w}_3 - \frac{\langle\vec{w}_3,\vec{v}_1\rangle}{\|\vec{v}_1\|^2}\vec{v}_1 - \frac{\langle\vec{w}_3,\vec{v}_2\rangle}{\|\vec{v}_2\|^2}\vec{v}_2$

For $k=1,2,\dots,n$

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

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

QR Decomposition

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

Then we get

$$\underbrace{\left(\vec{w}_1\middle|\vec{w}_2\middle|\cdots\middle|\vec{w}_n\right)}_{A \;\; m\times n} = \underbrace{\left(\vec{v}_1\middle|\vec{v}_2\middle|\cdots\middle|\vec{v}_n\right)}_{\text{orthogonal columns}}\underbrace{\begin{pmatrix}1 & \alpha_{12} & \alpha_{13} & \cdots\\ & 1 & \alpha_{23} & \cdots\\ & & 1 & \ddots\\ & & & 1\end{pmatrix}}_{n\times n}$$

$$\alpha_{ik} = \frac{\langle\vec{w}_k,\vec{v}_i\rangle}{\|\vec{v}_i\|^2}$$

The QR factorization is

$$A = QR$$

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

$Q = (\vec{u}_1|\vec{u}_2|\cdots|\vec{u}_n)$

$$R = \begin{pmatrix}\|\vec{v}_1\| & & 0\\ & \ddots & \\ 0 & & \|\vec{v}_n\|\end{pmatrix}\begin{pmatrix}1 & \alpha_{12} & \cdots & \alpha_{1n}\\ & 1 & & \vdots\\ & & \ddots & \alpha_{n-1,n}\\ 0 & & & 1\end{pmatrix}$$

Some remarks:

For $A\in\mathbb{F}^{m\times n}$ with LI columns we have $A = \underset{m\times n}{Q}\,\underset{n\times n}{R}$

i) $Q^HQ = I_n$

ii) R is upper triangular and invertible

iii) $A^HA = R^HR$ (Cholesky for A^HA)

Unitary (Orthogonal) Matrices

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

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

Then

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

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

iii) QS is unitary (orthogonal)

iv) $\langle Q\vec{x},Q\vec{y}\rangle = \langle\vec{x},\vec{y}\rangle \;\; \forall\vec{x},\vec{y}\in\mathbb{C}^n$ ($\mathbb{R}^n$)

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

$$\mathrm{Proj}_W(\vec{v}) = \sum_{i=1}^{n} \frac{\langle \vec{v}, \vec{v}_i \rangle}{\|\vec{v}_i\|_2} \vec{v}_i$$

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

Remarks:

- $\vec{v} \in W \Leftrightarrow \vec{v} = \mathrm{Proj}_W(\vec{v})$
- $\vec{v}_k = \vec{w}_k - \sum_{i=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$

  $= \vec{w}_k - \mathrm{Proj}_{\mathrm{span}\{\vec{v}_1, \dots, \vec{v}_n\}}(\vec{w}_k)$

Proposition: Let W be a subspace of V and let $\vec{p} = \mathrm{Proj}_W(\vec{v})$ 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 $Q = (\vec{u}_1 | \cdots | \vec{u}_n)$

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

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

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

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

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

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

Least Squares Minimization

Let $A \in \mathbb{F}^{m \times n}$ 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

$$\min_{\vec{x} \in \mathbb{F}^n} \|A\vec{x} - \vec{b}\|_2^2$$

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

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

$$A\vec{x}^* = QQ^H \vec{b} \Leftrightarrow R\vec{x}^* = Q^H \vec{b}$$

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