Inner Product

Inner product space: if there is a map $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{F}$

s.t. (i) $\langle u+v, w\rangle = \langle u,w\rangle + \langle v,w\rangle, \forall u,v,w \in V$

(ii) $\langle \alpha u, v\rangle = \alpha\langle u,v\rangle$

(iii) $\langle u,v\rangle = \overline{\langle v,u\rangle}$ (iv) $\langle v,v\rangle \geq 0$, $\langle v,v\rangle = 0 \Leftrightarrow v=0$

Ex.

(1) $V=\mathbb{C}^n$ $v=\begin{bmatrix} v_1 \\ \vdots \\ v_n\end{bmatrix}$, $v^H = [\bar{v}_1, \bar{v}_2 \cdots \bar{v}_n]$

$\langle u,v\rangle = v^H u = [\bar{v}_1\ \bar{v}_2 \cdots \bar{v}_n]\begin{bmatrix} u_1 \\ \vdots \\ u_n\end{bmatrix}$

$= \sum_{i=1}^{n} \bar{v}_i u_i$

Verify this use 4 properties. to proof i.p.

$\langle v,v\rangle = \sum_{i=1}^{n} |v_i|^2$

(2) $V=\mathbb{R}^n$, $\langle u,v\rangle = v^T u$

(3) $V=\mathbb{C}^{m\times n}$ $\boxed{\langle u,v\rangle = \mathrm{Trace}(V^H U)}$

$V=[v_1,\cdots v_n]$ $v_i \in \mathbb{C}^m$

$V^H = \begin{bmatrix} v_1^H \\ \vdots \\ v_n^H\end{bmatrix} \in \mathbb{C}^{n\times m}$

Norm A vector space V over $\mathbb{R}/\mathbb{C}$ is said to be a $\boxed{\text{normed v.s}}$ if there is a map $\|\cdot\| : V \to \mathbb{R}^+ = \{x\in\mathbb{R}; x\geq 0\}$

Satisfying:

(1) $\|\alpha\cdot v\| = |\alpha|\,\|v\|$ $\forall \alpha\in\mathbb{R}$ or $\mathbb{C}$, $v\in V$

(2) $\|v\|\geq 0$, $\|v\|=0 \Leftrightarrow v=0$

(3) $\|u+v\| \leq \|u\|+\|v\|$ (Triangle inequality)

E.x.

i) $V=\mathbb{C}^n$, integer $p\geq 1$

$\|v\|_p := \left(\sum_{i=1}^{n}|v_i|^p\right)^{\frac{1}{p}} \to L_p$-norm

Norm induced by inner product

Define:

$\|v\|_{i.p.} = \sqrt{\langle v,v\rangle}$

Can be shown that

$\|\cdot\|_{i.p.} : V \to \mathbb{R}^+$ is a norm

$\|v\|_{i.p.} = \sqrt{\langle v,v\rangle} = \sqrt{\sum_{i=1}^{n}|v_i|^2}$ - L_2 norm of V.

Set A, B. if $A\subseteq B$, $B\subseteq A$

then $A=B$.

If $\underset{n\times n}{A}$ is full rank, rank$(A)=n$.

Then $A^H A$ rank $=n$

$A^H A$ is invertible.

Orthogonality

$\mathbb{F}$: $\mathbb{R}$ or $\mathbb{C}$

V: v.s over $\mathbb{F}$ has a inner product $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{F}$

A set of non-zero vectors $\{v_1, v_2, \cdots, v_k\} \subseteq V$ is said to be

Orthogonal if $\langle v_i, v_j\rangle = 0$, $\forall i,j = 1,2,\cdots,k$ $i\neq j$. (pairwise)

Implications

Suppose $\{v_1, v_2, \cdots, v_k\}$ are orthogonal, $V=\mathbb{C}^n$

1) Define $\underset{(n\times k)}{V} = [v_1, v_2, \cdots, v_k] \Rightarrow V^H V = \underbrace{\begin{bmatrix} \|v_1\|^2_{i.p} & & 0 \\ & \|v_2\|^2_{i.p} & \\ 0 & & \ddots\ \|v_k\|^2 \end{bmatrix}}_{k\times k}$

Matrix with orthogonal columns.

$V^H V = \begin{bmatrix} v_1^H \\ \vdots \\ v_k^H\end{bmatrix}[v_1, v_2 \cdots v_k]$ $v_i^H v_j = 0$ $\forall i\neq j$.

(2) Let V be a vector space with i.p $\langle\cdot,\cdot\rangle$.

if $\{v_1, \cdots v_k\}$ are orthogonal, then $\{v_1, \cdots, v_k\}$ are also linear ind

Proof: $\sum_{i=1}^{k} \alpha_i v_i = 0$: $\langle \sum_{i=1}^{k}\alpha_i v_i, v_1\rangle = \langle 0, v_1\rangle = 0$

$\sum_{i}^{k} \alpha_i \langle v_i, v_1\rangle = 0$

$\|v_1\|^2_{i.p}\alpha_i = 0 \Rightarrow \alpha_i = 0$

(3) If $\dim(V) = r$, and $\{v_1, v_2, \cdots, v_r\}$ are orthogonal vectors in V

then $\{v_1, \cdots, v_r\}$. is also a basis of V. Orthogonal basis.

We need a basis for $S=R(Q)$.

Say rank$(Q) = r = \dim(R(Q))$. There exists a set of r linearly independent columns of Q.

Without loss of generality, let them be the first "r" columns.

$Q = [q_1\ q_2\ q_3 \cdots q_r\ q_{r+1} \cdots q_m]$

$\tilde{Q}_r = [q_1 \cdots q_r]$.

$b = Ax$ is equivalent to $\boxed{\underbrace{\tilde{Q}_r^H y}_{b} = (\underbrace{\tilde{Q}_r^H \tilde{Q}_r}_{A})x}$.

$\boxed{\hat{x} = (\tilde{Q}_r^H\tilde{Q}_r)^{-1}\tilde{Q}_r^H y}$

$\boxed{\hat{y}_s = \hat{y}_{R(Q)} = \tilde{Q}_r x = \tilde{Q}_r(\tilde{Q}_r^H\tilde{Q}_r)^{-1}\tilde{Q}_r^H y}$

$\hat{y}_{R(Q)} = P_{R(Q)}\, y$. O.P. of vector y (onto $R(Q)$) is a linear function of y

$\mathbb{C}^r \to R(Q)$

Application of Diagonalizable Matrix.

Solving system of linear Differential Equations:

Consider scalar case:

$x(t)$ has the following dynamics

$\frac{dx(t)}{dt} = \lambda x(t)$, $x(0)=c$

$x(t) = c\,e^{\lambda t}$

Consider $x_1(t), x_2(t), \cdots, x_m(t)$

$\frac{dx_1(t)}{dt} = a_{11}x_1(t) + a_{12}x_2(t) + \cdots + a_{1m}x_m(t)$

$\vdots$

$\frac{dx_m(t)}{dt} = a_{m1}x_1(t) + a_{m2}x_2(t) + \cdots + a_{mm}x_m(t)$

$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_m(t)\end{bmatrix}$, $A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm}\end{bmatrix}$

$\boxed{\frac{d}{dt}x(t) = A\,x(t)}$, $x(0) = c \in \mathbb{C}^m$

Suppose A is diagonalizable

$AP = P\Lambda \Rightarrow A = P\Lambda P^{-1}$

$\frac{dx(t)}{dt} = P\Lambda \underbrace{P^{-1}x(t)}_{y(t)}$ $y(t) = P^{-1}x(t)$

$P^{-1}\frac{dx(t)}{dt} = \Lambda y(t) \Rightarrow \frac{d(P^{-1}x(t))}{dt} = \Lambda y(t)$

$\frac{d}{dt}y(t) = \Lambda y(t)$

$\frac{dy_1(t)}{dt} = \lambda_1 y_1(t)$

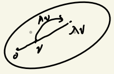
$\vdots$

$\frac{dy_m(t)}{dt} = \lambda_m y_m(t)$

$y(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_m e^{\lambda_m t}\end{bmatrix}$

$x(t) = P\begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_m e^{\lambda_m t}\end{pmatrix}$

Eigen Analysis



1. Consider $A\in\mathbb{C}^{m\times m}$ (complex square matrix).

A vector $v\neq 0$ is called an eigenvector of A

if $Av = \lambda v$, for some complex λ

λ: eigenvalue of A.

(λ, v) is eigen pair

if v is an eigenvector of A with eigenvalue λ.

then αv is also an eigenvector of A for the same λ.

$A(\alpha v) = \alpha\cdot Av$

$= \alpha\lambda v$

$= \lambda(\alpha v)$

Define, for a given λ.

$E_\lambda = \{v\in\mathbb{C}^n, v\neq 0, \text{s.t. } Av = \lambda v\}$.

= set of all possible eigenvectors corresponding to the same eigenvalue, λ.

Let $v_1, v_2 \in E_\lambda$. $A(\alpha v_1 + \beta v_2) = \alpha A v_1 + \beta A v_2$

$\alpha v_1 + \beta v_2 \in E_\lambda$. $= \alpha\lambda v_1 + \beta\lambda v_2$

$= \lambda(\alpha v_1 + \beta v_2)$.

Any linear combination of eigenvectors, for the same eigenvalue λ, is also a eigenvector.

$E_\lambda \cup \{0\}$ is subspace.

$v\in E_\lambda \Leftrightarrow Av = \lambda v$, $v\neq 0$.

$\Leftrightarrow (A-\lambda I)v = 0$, $v\neq 0$

$E_\lambda \cup\{0\} \Leftrightarrow v\in N(A-\lambda I)$

$\dim(N(A-\lambda I))$ = geometric multiplicity of λ, denoted by $g_A(\lambda)$

Nullity

$\boxed{g_A(\lambda)} = \dim(N(A-\lambda I))$

$N(A-\lambda I) \neq \{0\}$

$A-\lambda I$ is rank-deficient, $\Leftrightarrow \det(A-\lambda I) = 0$

rank$(A-\lambda I) < m$

$\det(A-\lambda I)$ is a polynomial in λ of degree m.

$\Rightarrow \det(A-\lambda I) = c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_m\lambda^m$

$= c(\lambda-\lambda_0)(\lambda-\lambda_1)\cdots(\lambda-\lambda_m)$

$\Leftrightarrow \lambda$ is a root of polynomial $\underline{c_0 + c_1\lambda + \cdots c_m\lambda^m}$.

characteristic polynomial of A.

$P_A(\lambda) = c_0\ c_1\lambda + \cdots c_m\lambda^m$

Suppose $P_A(\lambda)$ has k distinct roots, $\lambda_1 \sim \lambda_k$

Let $n_1, n_2, \cdots, n_k$ respectively be the # times, $\lambda_1 \sim \lambda_k$ repeat

n_i = # times λ_i repeats ($\lambda_i \in \{\lambda_1 \cdots \lambda_k\}$) (Algebraic Multiplicity of λ_i, denoted by $a_A(\lambda_i)$)

$\sum_{i=1}^{k} n_i = m$

Similar Matrices.

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

A, B are called Similar if there exists an invertible Matrix $P\in\mathbb{C}^{m\times m}$

s.t. $P^{-1}AP = B$. $A = PBP^{-1}$

Claim: A, B share same eigenvalues

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

$= \det(PBP^{-1} - \lambda PP^{-1})$

$= \det(P(B-\lambda I)P^{-1})$ rule of det.

$= \det(P)\det(B-\lambda I)\det(P^{-1})$

$= \det(P)\det(B-\lambda I)\frac{1}{\det(P)}$

$= \det(B-\lambda I)$

Orthogonal Projection

Given $y\in V$, $S\subseteq V$

$\min \|y-\hat{y}\|_{ip}$, s.t. $\hat{y}\in S$. — (OP)

Suppose a vector $\hat{y}_s$ exists in S. with the following property:

$\langle y-\hat{y}_s, v\rangle = 0$, $\forall v\in S$.

Property ①: Orthogonal principle:

$\hat{y}_s$ is called the orthogonal projection of y onto S

How to find $\hat{y}_s$?

① Find a basis of S $\{b_1 \ldots b_k\}$

② Solve $\boxed{b = A\hat{x}}$ $b\in\mathbb{C}^k$, $A\in\mathbb{C}^{k\times k}$ ($k=\dim(S)$) Normal Equation.

③ $\hat{y}_s = \sum_{i=1}^{k}\hat{x}_i b_i = [b_1\ b_2 \cdots b_k]\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_k\end{bmatrix}$

Application:

Let $V = \mathbb{C}^m$ $Q\in\mathbb{C}^{m\times n}$, $S = R(Q)$, $\langle\cdot,\cdot\rangle : \mathbb{C}^m\times\mathbb{C}^m \to \mathbb{C}$

$S\in\mathbb{C}^m$ $\langle u,v\rangle = v^H u$

$\|v\|_{ip} = \|v\|_2^2 = \sqrt{\sum_{i=1}^{m}|v_i|^2}$

Given $y\in\mathbb{C}^m$; solve $\min_{x\in\mathbb{C}^n}\|y - Qx\|_2$, $x\in\mathbb{C}^n$

$= \min_{\hat{y}\in\mathbb{C}^m}\|y-\hat{y}\|_2$

$\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}$ 3x4 matrix

$b = \begin{bmatrix} \langle b, v_1\rangle \\ \langle b, v_2\rangle \\ \vdots \\ \langle b, v_k\rangle \end{bmatrix}$ $(k\times 1)$ $\underset{(k\times k)}{A} = \begin{bmatrix} \langle v_1, v_1\rangle & \langle v_2, v_1\rangle & \cdots & \langle v_k, v_1\rangle \\ \langle v_1, v_2\rangle & \langle v_2, v_2\rangle & \cdots & \langle v_k, v_2\rangle \\ \vdots & & \ddots & \vdots \\ \langle v_1, v_k\rangle & \langle v_2, v_k\rangle & \cdots & \langle v_k, v_k\rangle \end{bmatrix}$ $x = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$ $k\times 1$

Diagonalizable Matrix

A matrix $A\in\mathbb{C}^{m\times m}$ is diagonalizable if A is similar to a diagonal matrix.

i.e. there exists invertible P s.t.

$P^{-1}AP = \begin{bmatrix} r_1 & & 0 \\ & r_2 & \\ 0 & & \ddots\ r_m \end{bmatrix}$

Implication:

Suppose A is diagonalizable there exists P, $r_1 \cdots r_m \in \mathbb{C}$

s.t. $P^{-1}AP = \begin{bmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_m\end{bmatrix}$ — $P = [P_1\ P_2 \cdots P_m]$

$\Rightarrow AP = P\begin{bmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_m\end{bmatrix}$

$\Leftrightarrow Ap_i = p_i r_i$, $i = 1 \ldots m$

p_i is an eigenvector of A and r_i is the corresponding eigenvalue

$\Leftrightarrow$ A has a set of m linearly independent eigen vectors. (given by columns of P)

A must be diagonalizable.

$Av_1 = \lambda_1 v_1$

$Av_m = \lambda_m v_m$.

$[Av_1\ Av_2 \cdots Av_m] = [v_1\ v_2 \cdots v_m]\begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots\ \lambda_m\end{bmatrix}$

$AV = V\Lambda \Rightarrow \boxed{V^{-1}AV = \Lambda}$.

Every $A\in\mathbb{C}^{m\times m}$ has eigenvalues & eigenvectors but it May Not Have a set of "m" linearly independent eigenvectors: Hence Not Every is diagonalizable

1) **Problem 1: Orthogonal Projection Matrices.** Let $\mathcal{M}$ and $\mathcal{N}$ be subspaces of $\mathbb{C}^n$, and consider the associated orthogonal projectors $\mathbf{P}_{\mathcal{M}}$ and $\mathbf{P}_{\mathcal{N}}$.

a) Prove that $\mathbf{P}_{\mathcal{M}}\mathbf{P}_{\mathcal{N}} = \mathbf{0}$ if and only if $\mathcal{M} \perp \mathcal{N}$.
b) Is it true that $\mathbf{P}_{\mathcal{M}}\mathbf{P}_{\mathcal{N}} = \mathbf{0}$ if and only if $\mathbf{P}_{\mathcal{N}}\mathbf{P}_{\mathcal{M}} = \mathbf{0}$? Justify
c) Show $R(\mathbf{P}_{\mathcal{M}} + \mathbf{P}_{\mathcal{N}}) = R(\mathbf{P}_{\mathcal{M}}) + R(\mathbf{P}_{\mathcal{N}})$.

Hw3

$\boxed{\text{Recall Hw3, projection matrix (projector) is symmetric matrix where } P^T = P \in R^{n\times n},\ P^2 = P}$

a). $P_M P_N = 0 \Rightarrow P_M^T P_n = 0$

$\Leftrightarrow \langle P_m, P_n \rangle = 0 \quad P_m = [p_1, p_2, \cdots, p_n]\ P_N = [P_1, \ldots, P_m]$,

$\langle p_j, p_i \rangle = 0 \ \forall P_{im}, P_{jn}$.

$\Leftrightarrow R(P_n) \perp R(P_m)$

$\Leftrightarrow M \perp N$

b). $P_M P_N = 0 \Leftrightarrow P_M^T P_N^T = 0 \Leftrightarrow (P_M P_N)^T \Leftrightarrow P_N^T P_M^T \Leftrightarrow P_N P_M = 0$.

c). $[P_M\ P_N][P_M\ P_N]^T = P_M \cdot P_M + P_N \cdot P_N$
$= P_M^2 + P_N^2$
$= P_M + P_N$

$R(P_M + P_N) = R([P_M\ P_N][P_M\ P_N]^T)$
$= R([P_M\ P_N])$
$= R(P_M) + R(P_N)$

2) **Problem 2: Orthonormal Basis Expansion and Parseval's Theorem.** Suppose we are given a set of orthonormal vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\}$ of an inner product vector space $\mathcal{U}$

a) Let $\mathbf{x} \in \mathcal{U}$, we can find a unique representation of $\mathbf{x} = \sum_{i=1}^N \alpha_i \mathbf{u}_i$. Prove

$$||\mathbf{x}||_2^2 = \sum_{i=1}^N |\alpha_i|^2 \tag{1}$$

(Note: This is known as Parseval's identity)

b) Suppose you have a subset of orthonormal vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_s\}$ (where $s < N$) from the given basis. Show that any vector $\mathbf{v} \in \mathcal{U}$ satisfies

$$||\mathbf{v}||_2^2 \geq \sum_{i=1}^s |\langle \mathbf{v}, \mathbf{u}_i \rangle|^2 \tag{2}$$

Lec7.

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

a). $\langle u_i, u_j \rangle = 1,\ \forall i = j$

$\langle u_i, u_j \rangle = 0$, when $i \neq j$.

$||x||_2^2 = \langle x, x \rangle = \langle \sum_{i=1}^N \alpha_i u_i, \sum_{i=1}^N \alpha_i u_i \rangle$

$= \sum_{i=1}^N \langle \alpha_i u_i, \sum_{i=1}^N \alpha_i u_i \rangle$

$= \sum_{i=1}^N \alpha_i \langle u_i, \sum_{i=1}^N \alpha_i u_i \rangle$

$= \sum_{i=1}^N \alpha_i \sum_{i=1}^N \bar{\alpha}_i \langle u_i, u_i \rangle$

$\forall i = i$

$= \sum_{i=1}^N |\alpha_i|^2$

b). $||v||_2^2 = \sum_{i=1}^N |\alpha_i|^2$

$\alpha_i = \dfrac{\langle v, u_i \rangle}{||u_i||_2^2} = \langle v, u_i \rangle$

$\sum_{i=1}^s |\langle v, u_i \rangle|^2 = \sum_{i=1}^s |\alpha_i|^2$

$\sum_{i=1}^N |\alpha_i|^2 \geq \sum_{i=1}^s |\alpha_i|^2$ for $N > S$.

3) **Problem 3: Range Space perpendicular to Null Spaces.**
Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ satisfy $\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$. Show that $R(\mathbf{A}) \perp N(\mathbf{A})$, i.e, show that for all $\mathbf{x} \in R(\mathbf{A})$, $\mathbf{y} \in N(\mathbf{A})$, $\mathbf{x}^H\mathbf{y} = 0$

DiS.

$N(A) = N(A^H A)$

let $x \in R(AA^H) \Rightarrow x = AA^H b = A(A^H b) \therefore x \in R(A)$.

$\therefore R(AA^H) \subset R(A)$.

$\because \dim(R(A)) = m - \dim(N(A)) = m - \dim(N(AA^H))$
$= \dim(R(AA^H))$

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

let $x \in R(A)$ $y \in N(A)$. $y \in N(AA^H)$

$x = AA^H z$, $Ay = 0$. $AA^H y = 0$

$x^H y = (AA^H z)^H y = z^H A^H A y = z(AA^H y) = 0$.

$\therefore R(A) \perp N(A)$

4) **Problem 4: Householder Reflections.** A Householder matrix is defined as

$$\mathbf{Q} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^\mathbf{T}$$

for a unit vector $\mathbf{u} \in \mathbb{R}^n$
a) Show that Q is orthogonal.
b) Show that $\mathbf{Qu} = -\mathbf{u}$ and that $\mathbf{Qv} = \mathbf{v}$ for every $\mathbf{v} \perp \mathbf{u}$. Thus, the linear transformation $\mathbf{y} = \mathbf{Qx}$ reflects $\mathbf{x}$ through the hyperplane with normal vector $\mathbf{u}$.
c) Given $\mathbf{y}$, find $\mathbf{x}$ such that $\mathbf{y} = \mathbf{Qx}$.
d) Given nonzero vectors $\mathbf{x}$ and $\mathbf{y}$, find a unit vector $\mathbf{u}$ such that $(\mathbf{I} - \mathbf{2uu^T})\mathbf{x} \in \text{span}(\mathbf{y})$, in terms of $\mathbf{x}$ and $\mathbf{y}$.

$u \in R^n$ $Q \in R^{n\times n}$ If Q is orthogonal $Q^TQ = QQ^T = I$

(a). $QQ^T = (I - 2uu^T)(I - 2uu^T)^T \quad (I - 2uu^T)^T = I - 2uu^T$

$= (I - 2uu^T)(I - 2uu^T)$

$= I - 2uu^T - 2uu^T + 4u\underbrace{u^Tu}_{I}u^T$

$= I$

Unitary matrix $u^Tu = I$.

$Q^TQ = QQ^T = I$

$\therefore$ Q is orthogonal.

(b). $Qu = (I - 2uu^T)u = u - 2uu^Tu = u - 2u = -u$.

$Qv = (I - 2uu^T)v = v - 2uu^Tv \quad \because v \perp u \quad \therefore u^Tv = 0$

$\therefore v - 0 = v$

$\because v \perp u = 0 \quad v = x - (u^Tx)u \ \forall x \in R^n \quad \because$ any vector can be projected on any subspaces.

$\therefore Qx = -(u^Tx)u + v$

Interpreted as the reflection of x through a hyperplane with normal vector u.

c) Since Q is symetric and orthogonal where $x = Q^{-1}y = Qy$ which is reflection back from y.

(4)

(d) The question is asking to find the vector $\mathbf{u}$ such that $\mathbf{x}$ and $\alpha\mathbf{y}$ for some α are reflections of each other through the hyperplane $\{\mathbf{z} : \mathbf{u}^T\mathbf{z} = 0\}$. Since $\mathbf{Qx}$ is given to be in the span of $\mathbf{y}$ we can write $\mathbf{Qx} = \alpha\mathbf{y}$ for some constant α, or $\|\mathbf{x}\| = \|\mathbf{Qx}\| = \|\alpha\mathbf{y}\|$, which implies that $\alpha = \pm\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|}$.
We consider the case $\mathbf{Qx} = (\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|})\mathbf{y} = \bar{\mathbf{x}}$, but the other case works equally well and provides an alternative answer. Now note that if the hyperplane $\{\mathbf{z} : \mathbf{u}^T\mathbf{z} = 0\}$ reflects $\mathbf{x}$ to $\bar{\mathbf{x}}$ and vice versa, then $\mathbf{x} - \bar{\mathbf{x}}$ is normal to the hyperplane. To see this $\mathbf{x} - \bar{\mathbf{x}} = (\mathbf{I} - \mathbf{Q})\mathbf{x} = 2\mathbf{u^Txu}$.
Suppose $\mathbf{x} \notin \mathbf{span}(\mathbf{y})$ so $\mathbf{x} - \bar{\mathbf{x}} \neq 0$(If that is not the case, then we can not represent $\mathbf{u}$ in term of $\mathbf{x}$ and $\mathbf{y}$). So:
$\mathbf{u} = \pm\frac{\mathbf{x}-\bar{\mathbf{x}}}{\|\mathbf{x}-\bar{\mathbf{x}}\|} = \pm\frac{\mathbf{x}-(\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|})\mathbf{y}}{\|\mathbf{x}-(\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|})\mathbf{y}\|} = \pm\frac{\|\mathbf{y}\|\mathbf{x}-\|\mathbf{x}\|\mathbf{y}}{\|\|\mathbf{y}\|\mathbf{x}-\|\mathbf{x}\|\mathbf{y}\|}$
which is the desired unit normal vector.
Alternatively, if we take $-\mathbf{x}$ as the reflection of $\mathbf{x}$, then $\mathbf{u} = \pm\frac{\mathbf{x}+\bar{\mathbf{x}}}{\|\mathbf{x}+\bar{\mathbf{x}}\|} = \pm\frac{\|\mathbf{y}\|\mathbf{x}+\|\mathbf{x}\|\mathbf{y}}{\|\|\mathbf{y}\|\mathbf{x}+\|\mathbf{x}\|\mathbf{y}\|}$.

5) **Problem 5: System Identification.** Consider a system whose input $x(n)$ and output $y(n)$ are related by:

$$y(n) = \sum_{k=0}^{L-1} h(k)x(n-k), \quad n = 0, 1, 2, \ldots \tag{3}$$

Here $h(n)$ is called the impulse response of the system. Suppose you are given an input signal $\overline{x}(n)$ (non-zero for all n) and are able to observe a noisy version ($\overline{y}(n)$) of the output of the system, contaminated with noise $w(n)$, i.e., you observe

$$\overline{y}(n) = \sum_{k=0}^{L-1} h(k)\overline{x}(n-k) + w(n), \quad n = 0, 1, 2, \ldots \tag{4}$$

Using the idea of orthogonal projection, describe a method to estimate the impulse response $h(n)$ using $\overline{y}(n)$ and $\overline{x}(n)$. In the absence of noise, under what conditions can you exactly identify $h(n)$? Justify your answer.

$$\min_{h \in R^L} \sum_{n=n_1}^{n_2} \left(\bar{y}(n) - \sum_{k=0}^{L-1} h(k)\bar{x}(n-k)\right)^2$$

$$= \min_{h \in R^L} \left\| \begin{bmatrix} \bar{y}(n_1) \\ \vdots \\ \bar{y}(n_2) \end{bmatrix} - \begin{bmatrix} \bar{x}(n_1) & \cdots & \bar{x}(n_1 - L + 1) \\ \vdots & & \\ \bar{x}(n_2) & \cdots & \bar{x}(n_2 - L + 1) \end{bmatrix} \begin{bmatrix} \bar{h}(0) \\ \vdots \\ \bar{h}(L-1) \end{bmatrix} \right\|_2^2$$

$$= \min_{h \in R^L} ||\bar{y} - \bar{X}h||_2^2.$$

$h = (\bar{x}^T\bar{x})\bar{x}^T\bar{y}$

6) **Problem 6: Variations of Orthogonal Projection**
a) Given $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, consider the following problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x} + \mathbf{c}\|_2 \tag{5}$$
$$s.t.\ \mathbf{Ax} = \mathbf{b} \tag{6}$$

Cast it as an orthogonal projection problem. Identify the subspace you are projecting on? What is the point being projected?

b) Given $\mathbf{x}_0 \in \mathbb{R}^n$, $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$, solve the following problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}_0 - \mathbf{x}\|_2 \tag{7}$$
$$s.t\ \mathbf{a}^T\mathbf{x} = b \tag{8}$$

Derive the solution in closed form.

(a). If $b \notin R(A)$, $Ax = b$ does not have solution.

If $b \in R(A)$, rank$(A) = n$ $\quad BA = I$

$x = Bb$

$\min_{x \in R^n} ||x + c||_2 = \min_{b \in R^m} ||Bb + c||_2$

This problem is finding a point x in set $\{x|Ax=b\}$, projecting onto a subspace $R(A) \in R^n$

If $b \in R(A)$ rank$(A) < n$

$Ax = b \Leftrightarrow x = x_0 + y \quad y \in N(A)$.

problem is $\min ||y + x_0 + c||_2$ st $y \in N(A)$

least square projection of $-(x_0 + c)$ onto $N(A)$

(b).

Derive a closed form of the solution.
Hints: We need to solve the orthonormal projection for $N(\mathbf{a}^T)$ first. As already shown in Homework 3, $N(\mathbf{A}^T) \perp R(\mathbf{A})$ holds for any matrix $\mathbf{A} \in \mathbb{R}^{m\times n} \implies N(\mathbf{a}^T) \perp R(\mathbf{a})$.
Clearly, orthonormal basis for $R(\mathbf{a})$ can be represented as follows along with its corresponding orthogonal projection matrix

$$\mathbf{u_a} = \frac{\mathbf{a}}{||\mathbf{a}||_2}, \text{ with corresponding projection matrix } \mathbf{u_a u_a^T} = \frac{\mathbf{aa^T}}{||\mathbf{a}||_2^2} \tag{40}$$

Thus,

$$\text{the orthonormal projector onto } N(\mathbf{a}^T) = \mathbf{I} - \frac{\mathbf{aa^T}}{||\mathbf{a}||_2^2} \tag{41}$$

Please make sure to complete the remaining steps by yourself.

3. Matrix Rank. Show the following identities about rank.

(a) If $\mathbf{A} \in \mathbb{F}^{m\times n}$, $\mathbf{B} \in \mathbb{F}^{n\times k}$, then
$$\text{rank}(\mathbf{B}) \leq \text{rank}(\mathbf{AB}) + \dim(\mathcal{N}(\mathbf{A}))$$
(b) If $\mathbf{A} \in \mathbb{F}^{m\times n}$, $\mathbf{B} \in \mathbb{F}^{m\times n}$, then
$$\text{rank}(\mathbf{A} + \mathbf{B}) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$$
(c) Suppose $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m\times m}$. Then, show that if $\mathbf{AB} = \mathbf{0}$, then
$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \leq m$$
(d) Suppose $\mathbf{A} \in \mathbb{F}^{m\times m}$. Then, show that $\mathbf{A}^2 = \mathbf{A}$ if and only if
$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{A} - \mathbf{I}) = m$$

Solution:

(a) $\mathbf{A} \in \mathbb{F}^{m\times n}$, $\mathbf{B} \in \mathbb{F}^{n\times k}$
The range space of $\mathbf{B}$ denoted $\mathcal{R}(\mathbf{B})$ is a subspace of $\mathbb{F}^n$.
Restrict the linear map $\mathbf{A}$ to $\mathcal{R}(\mathbf{B})$ i.e. consider new linear map $f : \mathcal{R}(\mathbf{B}) \to \mathbb{F}^m$ given by $f(\mathbf{x}) = \mathbf{Ax}$ for all vectors $\mathbf{x}$ in $\mathcal{R}(\mathbf{B})$.
Domain of f is $\mathcal{R}(\mathbf{B})$ and hence dimension of domain is $\text{rank}(\mathbf{B})$.

$$\mathcal{R}(f) = \{\mathbf{Ay} : \mathbf{y} \in \mathcal{R}(\mathbf{B})\} = \{\mathbf{ABx} : \mathbf{x} \in \mathbb{F}^k\} = \mathcal{R}(\mathbf{AB})$$
$$\mathcal{N}(f) = \{\mathbf{y} \in \mathcal{R}(\mathbf{B}) : \mathbf{Ay} = \mathbf{0}\} = \mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})$$

We can apply Rank-Nullity Theorem on this new map f.

$$\begin{aligned} \text{rank}(\mathbf{B}) &= \dim(\mathcal{R}(f)) + \dim(\mathcal{N}(f)) \\ &= \text{rank}(\mathbf{AB}) + \dim(\mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})) \\ &\leq \text{rank}(\mathbf{AB}) + \dim(\mathcal{N}(\mathbf{A})) \end{aligned}$$

(b) $\mathbf{A} \in \mathbb{F}^{m\times n}$, $\mathbf{B} \in \mathbb{F}^{m\times n}$

$$\begin{aligned} \mathcal{R}(\mathbf{A} + \mathbf{B}) &= \{(\mathbf{A} + \mathbf{B})\mathbf{x} : \mathbf{x} \in \mathbb{F}^n\} \\ &= \{\mathbf{Ax} + \mathbf{Bx} : \mathbf{x} \in \mathbb{F}^n\} \\ &\subseteq \{\mathbf{Ax} : \mathbf{x} \in \mathbb{F}^n\} + \{\mathbf{Bx} : \mathbf{x} \in \mathbb{F}^n\} \\ &= \mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B}) \end{aligned}$$

Hence $\text{rank}(\mathbf{A} + \mathbf{B}) \leq \dim(\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B}))$
Let $\mathcal{B}_1, \mathcal{B}_2$ be basis for $\mathcal{R}(\mathbf{A})$ and $\mathcal{R}(\mathbf{B})$ respectively. Then $\mathcal{B}_1 \cup \mathcal{B}_2$ spans $\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})$. Hence $\dim(\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$.
Hence $\text{rank}(\mathbf{A} + \mathbf{B}) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$

(c) Let $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m\times m}$. From part (a),

$$\begin{aligned} \text{rank}(\mathbf{B}) &\leq \text{rank}(\mathbf{AB}) + \dim(\mathcal{N}(\mathbf{A})) \\ &= \dim(\mathcal{N}(\mathbf{A})) \\ &= m - \text{rank}(\mathbf{A}) \quad [\because \text{Rank Nullity for } \mathbf{A}] \end{aligned}$$

Hence $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \leq m$

(d) Let $\mathbf{A} \in \mathbb{F}^{m\times m}$
Need to show $\mathbf{A}^2 = \mathbf{A} \iff \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{A} - \mathbf{I}) = m$

($\Longrightarrow$) Given $\mathbf{A}^2 = \mathbf{A}$
From part (b),

$$m = \text{rank}(\mathbf{I}) = \text{rank}(\mathbf{A} + (\mathbf{I} - \mathbf{A})) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{I} - \mathbf{A})$$

giving $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{I} - \mathbf{A}) \geq m$.
From part (c),
$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{I} - \mathbf{A}) \leq m$$
Hence $m = \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{I} - \mathbf{A})$.
As $\text{rank}(\mathbf{I} - \mathbf{A}) = \text{rank}(\mathbf{A} - \mathbf{I})$, we get $m = \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{A} - \mathbf{I})$

($\Longleftarrow$) Given $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{A} - \mathbf{I}) = m$
Let $\mathbf{x} \in \mathcal{N}(\mathbf{A})$, then $(\mathbf{A} - \mathbf{I})(-\mathbf{x}) = -\mathbf{Ax} + \mathbf{Ix} = \mathbf{x}$.
Hence $\mathcal{N}(\mathbf{A}) \subseteq \mathcal{R}(\mathbf{A} - \mathbf{I})$.
As $\dim(\mathcal{N}(\mathbf{A})) = m - \text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A} - \mathbf{I})$ we get $\mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{A} - \mathbf{I})$.

$$\mathcal{R}(\mathbf{A}(\mathbf{A} - \mathbf{I})) = \{\mathbf{A}(\mathbf{A} - \mathbf{I})\mathbf{x} : \mathbf{x} \in \mathbb{F}^m\} = \{\mathbf{Ay} : \mathbf{y} \in \mathcal{R}(\mathbf{A} - \mathbf{I})\} = \{0\}$$

Hence $\mathbf{A}(\mathbf{A} - \mathbf{I}) = \mathbf{0}$ i.e. $\mathbf{A}^2 = \mathbf{A}$.