ECE 6555, HW #1

Teo Wilkening
due 9/8/22

#1) Let $\vec{y} \in \mathbb{R}^m$ and $\underline{H} \in \mathbb{R}^{m \times n}$. $\vec{y} \in Im(\underline{H})$ s.t. $\underline{H}\vec{x}_0 = \vec{y}$. $\vec{x}_0$ is a solution. (at least one)

show that $\{\vec{x} : \underline{H}\vec{x} = \vec{y}\} = \vec{x}_0 + ker(\underline{H})$

① $ker(\underline{H}) = \{\vec{x} : \underline{H}\vec{x} = 0\}$, $\vec{x} \in \mathbb{R}^n$. let $\vec{x}_k \in ker(\underline{H})$

then take $\underline{H}(\vec{x}_0 + \vec{x}_k) = \underline{H}\vec{x}_0 + \cancel{\underline{H}\vec{x}_k}^{0} = \underline{H}\vec{x}_0 = \vec{y} \Rightarrow \boxed{\vec{x}_0 + ker(\underline{H}) \in \{\vec{x} : \underline{H}\vec{x} = \vec{y}\}}$

(part 1)

② show that $\{\vec{x} : \underline{H}\vec{x} = \vec{y}\} \subset \vec{x}_0 + ker(\underline{H})$

let $\tilde{x}_0$ be another solution to $\underline{H}\vec{x} = \vec{y}$. then, $\tilde{x}_0 = \vec{x}_0 + \tilde{x}_0 - \vec{x}_0$

$= \vec{x}_0 + \vec{v}$, $v = \tilde{x}_0 - x_0$

$\underline{H}\vec{v} = \underline{H}\tilde{x}_0 - \underline{H}\vec{x}_0 = \vec{y} - \vec{y} = 0$

$\Rightarrow \vec{v} \in ker(\underline{H})$

if $v \in ker(\underline{H})$ then this implies

that $\boxed{\tilde{\vec{x}}_0 \in \vec{x}_0 + ker(\underline{H})}$, $\underline{H}\tilde{\vec{x}}_0 = \vec{y}$

Thus by ① and ② we have shown that $\{\vec{x} : \underline{H}\vec{x} = \vec{y}\} = \vec{x}_0 + ker(\underline{H})$

#2) 2-1) $V \in \mathbb{R}^n$ show that $\|\vec{x}\| \triangleq \sqrt{\langle \vec{x}, \vec{x} \rangle}$ is indeed a norm.

Norm properties: · positive definiteness: $\forall \vec{x} \in V$, ~~$\forall \alpha \in \mathbb{R}$,~~ $\|\vec{x}\| > 0$ and $\|\vec{x}\| = 0$ iff $\vec{x} = 0$

· homogeneity: $\forall \vec{x} \in V, \forall \alpha \in \mathbb{R}$, $\|\alpha\vec{x}\| = |\alpha| \, \|\vec{x}\|$

· subadditivity: $\forall \vec{x}, \vec{y} \in V$, $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$

i) P.D. :

$\sqrt{\langle \vec{x}, \vec{x} \rangle} > 0$ iff $\langle \vec{x}, \vec{x} \rangle > 0$ $\boxed{\checkmark}$ → Inner product properties.

① symmetry : $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$

② bilinearity : $\langle \vec{x}, \lambda\vec{y} + \mu\vec{z} \rangle = \lambda\langle \vec{x}, \vec{y} \rangle + \mu\langle \vec{x}, \vec{z} \rangle$

③ P.D. : $\forall \vec{x} \neq 0$, $\langle \vec{x}, \vec{x} \rangle > 0$

ii) Homogeneity: $\alpha\vec{x}$

$\|\alpha\vec{x}\| = \sqrt{\langle \alpha\vec{x}, \alpha\vec{x} \rangle} = \sqrt{\alpha^2 \langle \vec{x}, \vec{x} \rangle}$

$= |\alpha| \sqrt{\langle \vec{x}, \vec{x} \rangle} = |\alpha| \, \|\vec{x}\|$ $\boxed{\checkmark}$

iii) Subadditivity:

$\|\vec{x} + \vec{y}\| \; ? \; \|\vec{x}\| + \|\vec{y}\| \Rightarrow ($~~$\sqrt{\langle x, \ldots}$~~ $\sqrt{\langle x+y, x+y \rangle} \; ? \; \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle}$

$\Rightarrow (\sqrt{\langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle} \; ? \; \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle})^2 \Rightarrow \cancel{\langle x, x \rangle} + 2\langle x, y \rangle + \cancel{\langle y, y \rangle} \; ? \; \cancel{\langle x, x \rangle} + \cancel{\langle y, y \rangle} + 2\sqrt{\langle x, x \rangle \langle y, y \rangle}$

$\Rightarrow (\langle x, y \rangle \; ? \; \sqrt{\langle x, x \rangle \langle y, y \rangle})^2 \Rightarrow \langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$ by Cauchy-Schwartz inequality

$\boxed{\checkmark}$ requires x, y to be real-valued.

→

#2-1 cont'd)

Thus since $||\vec{x}|| \triangleq \sqrt{\langle \vec{x}, \vec{x} \rangle}$ satisfies all of the properties of a norm, it is a norm. $\square$

#2-2) Complex #'s vs $\mathbb{R}$ $\quad$ $^{+}$ = transpose conjugate. i.e. $\vec{x}^{+} \triangleq (\vec{x}^{*})^{T}$

inner product now a sesquilinear form: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle^{*}$

i) show that $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} : (\vec{x}, \vec{y}) \mapsto \vec{x}^T \vec{y}^* = \vec{y}^{+} \vec{x}$ is a valid inner product

ii) Conclude that $||\vec{x}|| \triangleq \sqrt{\vec{x}^{+} \vec{x}}$ is a norm

i-1) Symmetry let $\vec{x}, \vec{y} \in \mathbb{C}^n$

then $\vec{x}^T \vec{y}^* = \sum_{i=1}^{n} x_i y_i^* = \sum_{i=1}^{n} y_i^* x_i = \vec{y}^{+} \vec{x}_i$

i-2) bilinearity let $\vec{z} \in \mathbb{C}^n$ as well

$\langle \vec{x}, \lambda \vec{y} + \mu \vec{z} \rangle = \vec{x}^T (\lambda \vec{y} + \mu \vec{z})^* = \sum_{i=1}^{n} x_i (\lambda y_i + \mu z_i)^*$

$= \sum_{i=1}^{n} \lambda x_i y_i^* + \mu x_i z_i^* = \lambda \vec{x}^T \vec{y}^* + \mu \vec{x}^T \vec{z}^* = \lambda \langle \vec{x}, \vec{y} \rangle + \mu \langle \vec{x}, \vec{z} \rangle$

i-3) **P.D.** let $\vec{x} \neq 0$

$\Rightarrow \langle \vec{x}, \vec{x} \rangle = \vec{x}^T \vec{x}^* = \sum_{i=1}^{n} x_i x_i^*$, $\quad x_i = a_i + i b_i$, $a_i, b_i \in \mathbb{R}$

$x_i x_i^* = (a_i + i b_i)(a_i - i b_i)$

$= a_i^2 + b_i^2 > 0$

$\Rightarrow \sum_{i=1}^{n} x_i x_i^* > 0$

Since all the properties of an inner product are satisfied, then $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{y}^*$ is a valid inner product. $\square$

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

ii) Since $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} : (\vec{x}, \vec{y}) \mapsto \vec{x}^T \vec{y}^* = \vec{y}^{+} \vec{x}$

is an inner product of $\mathbb{C}^n$, it follows from the proof in #2-1 on this homework set that $||\vec{x}|| \triangleq \sqrt{\vec{x}^{+} \vec{x}}$ is a norm

$\square$

#3-1 Orthogonal Complement

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

Show that $W^{\perp}$ is a vector subspace → definition of a vector subspace in lecture 1

let $\vec{x}, \vec{y} \in W^{\perp}$, let $\lambda, \mu \in \mathbb{R}$, then $\langle \lambda\vec{x} + \mu\vec{y}, \vec{w} \rangle = \langle \lambda\vec{x}, \vec{w} \rangle + \langle \mu\vec{y}, \vec{w} \rangle$

$= 0 + 0$

$\Rightarrow \lambda\vec{x} + \mu\vec{y} \in W^{\perp}$ Thus, $W^{\perp}$ is a vector subspace by the definition of a vector subspace $\square$

#3-2 Show that $W \oplus W^{\perp} = V$ (direct sum)

The definition of direct sum is: $V = W \oplus W^{\perp}$ iff $\forall \vec{v} \in V$ $\exists$ a unique $(\vec{w}, \vec{w}^{\perp}) \in W \times W^{\perp}$ s.t. $\vec{v} = \vec{w} + \vec{w}^{\perp}$

Thus there are three conditions we must satisfy:

a) $W \oplus W^{\perp} \subset V$

b) $V \subset W \oplus W^{\perp}$

c) $\vec{w} + \vec{w}^{\perp}$ is unique (i.e. $\vec{w}_1 + \vec{w}_1^{\perp} \neq \vec{w}_2 + \vec{w}_2^{\perp}$)

3-2a) $W \oplus W^{\perp} \subset V$

By the definition of an orthogonal complement,

$W \subset V$, $W^{\perp} \subset V$ $\Rightarrow$ W and $W^{\perp}$ are subspaces of V.

Thus any two vectors $\vec{w} \in W$ and $\vec{w}^{\perp} \in W^{\perp}$ must also be in the vector space V. ~~Therefore~~ $\vec{w} + \vec{w}^{\perp} = \vec{\xi}$, $\vec{\xi} \in V$

Therefore $W \oplus W^{\perp} \subset V$ $\square$

3-2b) $V \subset W \oplus W^{\perp}$

let $W = \mathrm{Im}(\underline{H})$, let $\vec{v} \in V$. Then $\vec{v} = P_H\vec{v} + (I - P_H)\vec{v} = \vec{y} + \vec{y}^{\perp}$.

($\underline{H}$ is full rank) ~~Thus~~ where $\vec{y} \in W$ and $\vec{y}^{\perp} \in W^{\perp}$ by the definition

($\underline{H}$ can always be a basis For W.) of the orthogonal projector P_H.

Thus, $\forall \vec{v} \in V$, $\exists$ $\vec{y} \in W$ and $\vec{y}^{\perp} \in W^{\perp}$ s.t. $\vec{v} = \vec{y} + \vec{y}^{\perp}$.

Therefore $V \subset W \oplus W^{\perp}$ $\square$

#3-2c) $\vec{w} + \vec{w}^{\perp} = \vec{v}$ unique

assume $\vec{w} + \vec{w}^{\perp} = \vec{v}$ is <u>not</u> unique, then the following would be ~~is~~ true:

$$\vec{w}_1 + \vec{w}_1^{\perp} = \vec{v} = \vec{w}_2 + \vec{w}_2^{\perp} \quad , \quad \vec{w}_1, \vec{w}_2 \in W \; ; \; \vec{w}_1^{\perp}, \vec{w}_2^{\perp} \in W^{\perp}, \; \vec{w}_1 \neq \vec{w}_2, \; \vec{w}_1^{\perp} \neq \vec{w}_2^{\perp}$$

$$\Rightarrow \vec{w}_1 - \vec{w}_2 + \vec{w}_1^{\perp} - \vec{w}_2^{\perp} = 0$$

Now, represent the vectors using bases:

let $\underline{A} = \{\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_i\}$ be an orthonormal basis for W

let $\underline{B} = \{\vec{\beta}_1, \vec{\beta}_2, \ldots, \vec{\beta}_j\}$ be an orthonormal basis for $W^{\perp}$

<u>note</u> that $\vec{\alpha}_i, \vec{\beta}_j$ are linearly independent b/c constituent vectors are orthogonal. i.e. $\langle \vec{\alpha}_i, \vec{\beta}_j \rangle = \alpha_{i1}\beta_{j1} + \alpha_{i2}(\beta_{j2}) + \ldots + \alpha_{in}\beta_{jn} = 0$

then, $\vec{w} = \sum^{i} \lambda_i \vec{\alpha}_i \; , \; \lambda_i \in \mathbb{R}$

$\vec{w}^{\perp} = \sum^{j} \gamma_j \vec{\beta}_j \; , \; \gamma_j \in \mathbb{R}$

$$\Rightarrow \vec{w}_1 - \vec{w}_2 + \vec{w}_1^{\perp} - \vec{w}_2^{\perp} = \sum^{i} \lambda_{1i} \vec{\alpha}_i - \sum \lambda_{2i} \vec{\alpha}_i + \sum^{j} \gamma_{1j} \vec{\beta}_j - \sum^{j} \gamma_{2j} \vec{\beta}_j = 0$$

$$= \sum^{i} (\lambda_{1i} - \lambda_{2i}) \vec{\alpha}_i + \sum^{j} (\gamma_{1i} - \gamma_{2j}) \vec{\beta}_j = 0$$

By linear independence of $\vec{\alpha}_i, \vec{\beta}_j$, the above can be true iff $\lambda_{1i} = \lambda_{2i}$ and $\gamma_{1j} = \gamma_{2j}$. Thus ~~for $\vec{w}_1 + \vec{w}_1^{\perp} = \vec{w}_2 + \vec{w}_2^{\perp}$ to b~~

But then $\vec{w}_1 = \vec{w}_2$ and $\vec{w}_1^{\perp} = \vec{w}_2^{\perp}$ and this contradicts our assumption.

Thus, $\vec{w}_1 + \vec{w}_1^{\perp} \neq \vec{w}_2 + \vec{w}_2^{\perp}$ and hence $\vec{w}_1 + \vec{w}^{\perp} = \vec{v}$ is <u>unique</u>. $\square$

#3-3) Show that $(W^{\perp})^{\perp} = W \quad , \; W = Im(H) = \{\vec{w}\}, \; \vec{w} \in W$

by ~~definition~~ definition: $W^{\perp} = \{\vec{w}^{\perp} \in V : \forall \vec{w} \in W, \langle \vec{w}^{\perp}, \vec{w} \rangle = 0\} \triangleq \{\vec{w}^{\perp}\}$

then, $(W^{\perp})^{\perp} = \{\vec{y} \in V : \forall \vec{w}^{\perp} \in W^{\perp}, \langle \vec{y}, \vec{w}^{\perp} \rangle = 0\} \triangleq \{\vec{y}\}$

$$\Rightarrow \langle \vec{w}^{\perp}, \vec{w} \rangle = 0 = \langle \vec{y}, \vec{w}^{\perp} \rangle = \langle \vec{w}^{\perp}, \vec{y} \rangle \Rightarrow \langle \vec{w}^{\perp}, \vec{w} \rangle = \langle \vec{w}^{\perp}, \vec{y} \rangle$$

$\Rightarrow \vec{w} = \vec{y}$. Thus, $\{\vec{w}\} = \{\vec{y}\}$, that is, $W = (W^{\perp})^{\perp}$ $\square$

#4) Let $\underline{H} \in \mathbb{R}^{m \times n}$, Show that $Im(H^T) = Im(H^T H)$

a) $Im(\underline{H}^T \underline{H}) \subset Im(\underline{H}^T)$:

$Im(\underline{H}^T \underline{H}) \Rightarrow \vec{y} = \underline{H}^T \underline{H} \vec{x}$ for some $\vec{x} \in \mathbb{R}^n$.

$Im(\underline{H}^T) \Rightarrow \vec{z} = \underline{H}^T \cancel{\vec{u}}$, $\vec{u} \in \mathbb{R}^m$

Let $\vec{w} = \underline{H} \vec{x} \Rightarrow \vec{w} \in \mathbb{R}^m$

Then $\vec{y} = \underline{H}^T \underline{H} \vec{x} = \underline{H}^T \vec{w}$ and $\underline{H}^T \vec{w} \in Im(\underline{H}^T)$

Therefore $Im(\underline{H}^T \underline{H}) \subset Im(\underline{H}^T)$

$\square$

b) $Im(H^T) \subset Im(\underline{H}^T \underline{H})$

~~$Im(\underline{H}^T) \subset Im(\underline{H}^T \underline{H}) \Leftrightarrow Im(\underline{H}^T)^\perp \subset Im(\underline{H}^T \underline{H})^\perp$~~

~~$\Leftrightarrow ker(\underline{H}^T) \subset ker(\underline{H}^T \underline{H})$~~

~~let $\vec{x} \in ker(\underline{H}^T)$ s.t. $\underline{H}^T \vec{x} = 0$~~

Let $\vec{y} \in Im(\underline{H}^T)$, ~~so~~ then $\exists\, \vec{u}$ s.t. $\vec{y} = \underline{H}^T \vec{u}$, $\vec{u} \in \mathbb{R}^m$

note: $\mathbb{R}^n = Im(\underline{H}) \oplus Im(\underline{H})^\perp$

note also: $Im(\underline{H})^\perp = ker(\underline{H}^T)$

then, $\vec{u} = \tilde{\vec{u}} + \hat{\vec{u}}$, $\tilde{\vec{u}} \in Im(\underline{H})$, $\hat{\vec{u}} \in ker(\underline{H}^T)$, $\tilde{\vec{u}} = H\vec{v}$, $\vec{v} \in \mathbb{R}^n$

and $\vec{y} = \underline{H}^T(\tilde{\vec{u}} + \hat{\vec{u}}) = \underline{H}^T(\underline{H}\vec{v} + \cancelto{0}{\hat{\vec{u}}}) = \underline{H}^T \underline{H} \vec{v}$

$\Rightarrow \exists\, \vec{v}$ s.t. $\vec{y} \in Im(\underline{H}^T \underline{H}) \Rightarrow Im(\underline{H}^T) \subset Im(\underline{H}^T \underline{H})$

$\square$

Therefore, $Im(\underline{H}^T \underline{H}) = Im(\underline{H}^T)$

$\square$

#3-2 cont'd)

$V = \text{span}(\{v_i\}^n_i) \in \mathbb{R}^n$ [redacted] $\mathbb{R}^n$ vector space

$= \text{span}($[redacted]$\{w_i^\perp\}^{n-k}_i)$ [redacted] $\mathbb{R}^n$ vector space

n [redacted] onormal basis v[redacted] $\in \mathbb{R}^n \Rightarrow \text{span}($[redacted]$) = \text{span}(\{w_i\}, \{w_i^\perp\})$ [redacted] vector space

use this? or s[redacted] else? ^ ... dimensions...

identically.

#5

$$J(x) \triangleq (\vec{y} - \underline{H}\vec{x})^T \underline{W} (\vec{y} - \underline{H}\vec{x}) \quad , \quad W > 0 \text{ \& symmetric}$$

i.e. $x^T W x > 0$ for $z \neq \{0\}, z \in \mathbb{R}^n$

Q1) Show $J(x) > 0$

Since W is positive definite, then by definition for a real vector $\vec{z} \in \mathbb{R}^n$, $\vec{z} \neq \{0\}$ $\vec{z}^T W \vec{z} > 0$. If we let $\vec{z} = (\vec{y} - \underline{H}\vec{x})$ then we have

$$\vec{z}^T W \vec{z} = (\vec{y} - \underline{H}\vec{x})^T W (\vec{y} - \underline{H}\vec{x}) > 0 \qquad \square$$

Q2) Show that cost is zero iff $\vec{y} = \underline{H}\vec{x}$:

if $\vec{y} = \underline{H}\vec{x}$ then $J(\vec{x}) = (\underline{H}\vec{x} - \underline{H}\vec{x})^T \underline{W} (\underline{H}\vec{x} - \underline{H}\vec{x}) = 0^T W 0 = 0$

if $J(x) = 0$ then by the definition of a positive definite matrix, this can only be true if $\vec{z} = (\vec{y} - \underline{H}\vec{x}) = 0 \Rightarrow \vec{y} = \underline{H}\vec{x}$

Thus, $J(\vec{x}) = 0$ iff $\vec{y} = \underline{H}\vec{x} \qquad \square$

Q3) Normal equation for modified cost: $\underline{H}^T \underline{W} \underline{H} \vec{x} = \underline{H}^T \underline{W} \vec{y}$

So, $J(\vec{x}) \triangleq (\vec{y} - \underline{H}\vec{x})^T \underline{W} (\vec{y} - \underline{H}\vec{x}) = (\vec{y} - \underline{H}\vec{x})^T (\underline{W}\vec{y} - \underline{W}\underline{H}\vec{x})$

$$= (\vec{y}^T - \vec{x}^T \underline{H}^T)(\underline{W}\vec{y} - \underline{W}\underline{H}\vec{x})$$

$$= \vec{y}^T \underline{W} \vec{y} - \vec{x}^T \underline{H}^T \underline{W} \vec{y} - \vec{y}^T \underline{W} \underline{H} \vec{x} + \vec{x}^T \underline{H}^T \underline{W} \underline{H} \vec{x}$$

$$(\vec{x}^T \underline{H}^T \underline{W} \vec{y})^T = \vec{y}^T \underline{W}^T \underline{H} \vec{x} = \vec{y}^T \underline{W} \underline{H} \vec{x} \rightarrow$$

b/c $\vec{x}^T \underline{H}^T \underline{W} \vec{y} \in \mathbb{R}$, $\underline{W} = \underline{W}^T$

note:

$$\frac{\partial \vec{x}^T A \vec{x}}{\partial \vec{x}} = (A+A^T)\vec{x}$$

$$\frac{\partial \vec{y}^T A \vec{x}}{\partial \vec{x}} = A^T \vec{y}$$

$$\frac{\partial \vec{x}^T A \vec{y}}{\partial \vec{x}} = A\vec{y}$$

,
$$J(\vec{x}) = \vec{y}\,\underline{W}\,\vec{y} - \vec{y}^T \underline{W}\,\underline{H}x - \vec{y}^T \underline{W}\,\underline{H}\vec{x} + \vec{x}^T \underline{H}^T \underline{W}\,\underline{H}\vec{x}$$

then,

$$\frac{\partial J}{\partial \vec{x}} = 0 - 2\,\underline{H}^T \underline{W}^T \vec{y} + (\underline{H}^T \underline{W}\,\underline{H} + \underline{H}^T \underline{W}\,\underline{H})\vec{x}$$

(W is symmetric)

(assuming variables are real-valued for gradient of $J(\vec{x})$)

$$\Rightarrow \frac{\partial J(\vec{x})}{\partial \vec{x}} = -2\,\underline{H}^T \underline{W}^T \vec{y} + 2\,\underline{H}^T \underline{W}\,\underline{H}\vec{x} = 0$$

to minimize

$$\Rightarrow \underline{H}^T \underline{W}\vec{y} = \underline{H}^T \underline{W}\,\underline{H}\vec{x}$$

$\square$

#6 Projectors

$P_H \triangleq \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T$ is the projector onto $Im(\underline{H})$

6-1) show $P_H = P_H^T$:

$$P_H^T = (\underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)^T = \underline{H}(\underline{H}^T\underline{H})^{-T}\underline{H}^T \; ; \; (\underline{H}^T\underline{H})^{-T} = ((\underline{H}^T\underline{H})^{-1})^{T} = (\underline{H}^T\underline{H})^{-1}$$

$$= \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T$$

→ note also:

$$= P_H \quad \square \qquad P_H^2 = P_H P_H = P_H$$

6-2) Show $I - P_H$ is a projector onto $Im(\underline{H})^\perp$:

Let ~~W~~ W be a linear subspace of $V \in \mathbb{R}^n$ and let $\hat{\vec{y}}$ be the orthogonal projection of $\vec{y} \in V$ onto W, where $W \equiv Im(\underline{H})$

$W \oplus W^\perp = V$

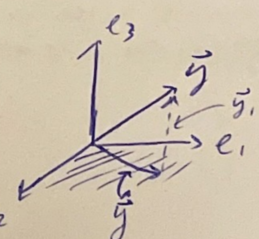
then $\hat{\vec{y}} = P_H \vec{y} = \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\vec{y}$

Now, $(I - P_H)\vec{y} = \vec{y} - P_H\vec{y} = \vec{y} - \hat{\vec{y}}$

Since $\mathbb{R}^n = Im(\underline{H}) \oplus Im(\underline{H})^\perp$, $\vec{y} = \vec{y}_0 + \vec{y}_1$ where $\vec{y}_0 \in W$, $\vec{y}_1 \in W^\perp$

Since $\hat{\vec{y}} = P_H\vec{y}$, it contains all of the components of $\vec{y}$ contained in W. Thus $\hat{\vec{y}} = \vec{y}_0$ and $\vec{y}_1 = \vec{y} - \hat{\vec{y}} \in W^\perp$

Graphically in $\mathbb{R}^3$:



Therefore, $\vec{y}_1 = (I - P_H)\vec{y}$ and $(I - P_H)$ is the projector ~~of $\vec{y}$~~ onto $Im(\underline{H})^\perp$

$\square$

#6-Q3 show that if $\vec{y} \in Im(\underline{H})$ then $P_H \vec{y} = \vec{y}$

$\Rightarrow \vec{y} = H\vec{x} \Rightarrow P_H \vec{y} = \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{H}\vec{x}$

$= \underline{H}\,\underline{I}\,\vec{x}$

$= \underline{H}\vec{x} = \vec{y}$

$\square$

注 P_H presumes $\underline{H}$ full rank.

$\Rightarrow$ $\underline{H}^T\underline{H}$ = Gramian Matrix $\in \mathbb{R}^{n \times n}$

$\Rightarrow$ by properties of Gramian Matrix $(\underline{H}^T H)$ is non-singular for $\underline{H}$ full rank

and $(\underline{H}^T\underline{H})^{-1}\underline{H}^T H = \underline{I}$

#7 Projection w/ Mahalanobis distance

Show $P_H = \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T$ should be replaced by $P_H = \underline{H}(\underline{H}^T\underline{W}\underline{H})^{-1}\underline{H}^T\underline{W}$

when $J(\vec{x}) = (\vec{y} - \underline{H}\vec{x})^T\underline{W}(\vec{y} - \underline{H}\vec{x})$

Recall from Q5-3 that the normal equation is:

$\underline{H}^T\underline{W}\vec{y} = \underline{H}^T\underline{W}\,\underline{H}\vec{x} \Rightarrow \vec{x}_0 = (\underline{H}^T\underline{W}\underline{H})^{-1}\underline{H}^T\underline{W}\vec{y}$ (best estimate of $\vec{x}$)

$\Rightarrow \hat{\vec{y}} = \underline{H}\vec{x}_0 = \underbrace{\underline{H}(\underline{H}^T\underline{W}\underline{H})^{-1}\underline{H}^T\underline{W}}_{P_H}\vec{y}$ (best estimate or "projection" of $\vec{y}$ onto solution space W = ~~span(H)~~ $Im(H)$)

$\Rightarrow P_H = \underline{H}(\underline{H}^T\underline{W}\underline{H})^{-1}\underline{H}^T\underline{W}$

$\square$

From solution: the projection is for what space??

$\Rightarrow$ define new inner prod: $\langle y, x \rangle = y^T W x$

For some $z \in Im(H)$ $\Rightarrow \langle y - \hat{y}, z \rangle = (y - \hat{y})^T W x = (y^T - (H(H^TWH)^{-1}H^TWy)^T) WHx$

$\Rightarrow = y^T(I - WH(H^TWH)^{-1}H^T)WHx = 0$

$\Rightarrow$ thus $y - \hat{y}$ is $\perp$ to $Im(H)$

$\Rightarrow \hat{y}$ is indeed the orthogonal projection.

#8 Constrained Least Square

$\min_x \|\vec{y} - \underline{H}\vec{x}\|_2^2$ subject to $\underline{A}\vec{x} = \vec{b}$. Equivalent to minimization of

unconstrained $\min_x \left(\|\vec{y} - \underline{H}\vec{x}\|_2^2 + \vec{\lambda}^T(\underline{A}\vec{x} - \vec{b}) \right)$, $\vec{\lambda}$ = Lagrange multipliers

allow us to perform minimization/max. for a constrained problem analytically (converts constraint into a cost)

$\longrightarrow$

(behind page 1) (physical paper)

8
~~12~~)

#8 cont'd)

note: $\vec{x}_0 = \vec{x}_{LS} = (\underline{H}^T\underline{H})^{-1}\underline{H}^T\vec{y}$, $\vec{y} = \underline{H}\vec{x}$

let $J(\vec{x}) = \|\vec{y} - \underline{H}\vec{x}\|_2^2 + \vec{\lambda}^T(\underline{A}\vec{x} - \vec{b})$

$= (\vec{y} - \underline{H}\vec{x})^T(\vec{y} - \underline{H}\vec{x}) + \vec{\lambda}^T(\underline{A}\vec{x} - \vec{b})$

min $J(\vec{x})$ is where $\nabla_{x,\lambda} J(\vec{x}, \vec{\lambda}) = \vec{0}$

$\Rightarrow \frac{\partial J}{\partial \vec{x}} = 2\underline{H}^T\underline{H}\vec{x} - 2\underline{H}^T\vec{y} + \underline{A}^T\vec{\lambda} = 0$

note: $\frac{\partial \vec{x}^T\underline{A}\vec{x}}{\partial \vec{x}} = (\underline{A} + \underline{A}^T)\vec{x}$

$\frac{\partial J}{\partial \vec{\lambda}} = \underline{A}\vec{x} - \vec{b} = 0 \Rightarrow \vec{x} = \underline{A}^{-1}\vec{b}$

$\frac{\partial \vec{y}\underline{A}\vec{x}}{\partial \vec{x}} = \underline{A}^T\vec{y}$

(assume $\underline{A}$ non-singular)

$\frac{\partial \vec{x}^T\underline{A}\vec{y}}{\partial \vec{x}} = \underline{A}\vec{y}$

$\Rightarrow$ solve for $\vec{\lambda}$:

$\Rightarrow \frac{\partial J}{\partial \vec{x}} = 2\underline{H}^T\underline{H}A^{-1}\vec{b} - 2\underline{H}^T\vec{y} + \underline{A}^T\vec{\lambda} = 0$

note: $\underline{I} = \underline{H}^T\underline{H}A^{-1}A(\underline{H}^T\underline{H})^{-1}$

$\Rightarrow$ ~~$0 = 2\underline{I}\,\underline{H}^T\underline{H}$~~

$\underline{A}^T\vec{\lambda} = -2\underline{H}^T\underline{H}A^{-1}\vec{b} + 2\underline{I}\underline{H}^T\vec{y}$

$= -2\underline{H}^T\underline{H}A^{-1}\vec{b} + 2\underline{H}^T\underline{H}A^{-1}\underline{A}\underbrace{(\underline{H}^T\underline{H})^{-1}\underline{H}^T\vec{y}}_{\vec{x}_{LS}}$

$\underline{A}^T\vec{\lambda} = 2[\underline{H}^T\underline{H}A^{-1}](\underline{A}\vec{x}_{LS} - \vec{b})$

$\rightarrow \vec{\lambda} = 2[A^{-T}\underline{H}^T\underline{H}A^{-1}](\underline{A}\vec{x}_{LS} - \vec{b})$

$= 2[\underline{A}(\underline{H}^T\underline{H})^{-1}A^T]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})$

$\square$

note: $AB^{-1}A = BA^{-1}B$

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

$\Rightarrow [A(\underline{H}^T\underline{H})^{-1}A^T]^{-1}$

$= A^{-T}(\underline{H}^T\underline{H})A^{-1}$

(#8-2 see last page)

#8 con't'd

8-2) Conclude that the constrained LS solution is:

$$\vec{x}_c = \vec{x}_{LS} - (\underline{H}^T\underline{H})^{-1}\underline{A}^T\left[\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T\right]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})$$

So, $\frac{\partial J(\vec{x},\vec{\lambda})}{\partial \vec{x}} = 2\underline{H}^T\underline{H}\vec{x} - 2\underline{H}^T\vec{y} + \underline{A}^T\vec{\lambda} = 0$

$$= \cancel{2}\underline{H}^T\underline{H}\vec{x} - \cancel{2}\underline{H}^T\vec{y} + \underline{A}^T\left(\cancel{2}\left[\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T\right]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})\right) = 0$$

$$= \underline{H}^T\underline{H}\vec{x} - \underline{H}^T\vec{y} + \underline{A}^T\left[\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T\right]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})$$

$\Rightarrow$ normal equation:

$$\underline{H}^T\underline{H}\vec{x} = \underline{H}^T\vec{y} + \underline{A}^T\left[\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T\right]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})$$

$\Rightarrow$ solution to normal equations:

$$\vec{x}_c = \underbrace{(\underline{H}^T\underline{H})^{-1}\underline{H}^T\vec{y}}_{\vec{x}_{LS}} + (\underline{H}^T\underline{H})^{-1}\underline{A}^T\underbrace{\left[\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T\right]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})}_{\vec{\lambda}/2}$$

$$\boxed{\vec{x}_c = \vec{x}_{LS} + (\underline{H}^T\underline{H})^{-1}\underline{A}^T\left[\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T\right]^{-1}(\underline{A}\vec{x}_{LS} - \vec{b})}$$

□

10/15)