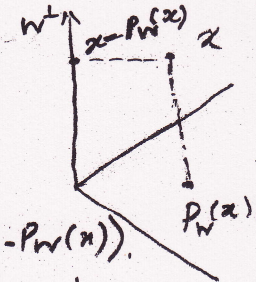
# Lecture 14

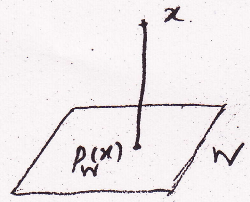



**Theorem:** (Projection Thm.) Let $W$ be a subspace of a f.d.i.p.s. $V$.
Then $V = W \oplus W^{\perp}$ i.e., $V = W + W^{\perp}$ and $W \cap W^{\perp} = \{0\}$.

**Proof:** Any $x \in V$ can be written as $x = P_W(x) + (x - P_W(x))$.
Therefore $x \in W + W^{\perp}$. This implies that $V \subseteq W + W^{\perp}$.
The other inclusion is obvious.

**Remarks:** (i) Note that $P_W(x)$ exists and it is unique
(ii) If $V = W \oplus W^{\perp}$ and $v \in V$ then there exist unique $w_1 \in W$ and $w_2 \in W^{\perp}$ such that $v = w_1 + w_2$. (why?).

**Nearest point:** If $W$ is the $xy$ plane in $\mathbb{R}^3$ and $x \in \mathbb{R}^3$, then we know that the orthogonal projection $P_W(x)$ is the nearest point for $x$ from $W$. We will see that this is true in general.




**Theorem:** Let $W$ be a subspace of a f.d.i.p.s. $V$ & let $x \in V$. Then $\|x - P_W(x)\| \le \|x - w\| \ \forall w \in W$.

**Proof:** Note that

$$\|x-w\|^2 = \|x - P_W(x) + P_W(x) - w\|^2 = \|x - P_W(x)\|^2 + \|P_W(x) - w\|^2 \text{ (why?)}$$
$$\geq \|x - P_W(x)\|^2.$$

**Four Fundamental Subspaces:**

Let $A = (a_{ij})_{m\times n}$ be an $m\times n$ matrix & consider $A : \mathbb{R}^n \to \mathbb{R}^m$.
We will use the following notation:

- $N(A)$ - Null space of $A$ (which in $\mathbb{R}^n$, i.e., in the domain of $A$)
- $C(A)$ - Column space of $A$, generated by column vectors,
  - which is nothing but the range of $A$
- $\bar{\mathcal{R}}(A)$ - row space of $A$, generated by row vectors.
  - which is a subspace of $\mathbb{R}^n$, the domain of $A$.

So, $N(A)$ & $\bar{\mathcal{R}}(A) \subseteq \mathbb{R}^n$ and $C(A) \subseteq \mathbb{R}^m$.

Since $A^T = (a_{ji})_{n\times m}$, $A^T : \mathbb{R}^m \to \mathbb{R}^n$ and hence $N(A^T) \subseteq \mathbb{R}^m$.

Note that $N(A)$ and $\bar{\mathcal{R}}(A)$ are subspaces of $\mathbb{R}^n$ – the domain of $A$ & $C(A)$ & $N(A^T)$ are subspaces of $\mathbb{R}^m$ = the range of $A$ & $\bar{\mathcal{R}}(A^T) = C(A)$.

**Relations between $N(A)$ and $\bar{\mathcal{R}}(A)$ (resp. $C(A)$ & $N(A^T)$):**

**Lemma:** $N(A) \perp \bar{\mathcal{R}}(A)$ in $\mathbb{R}^n$ and $N(A^T) \perp C(A)$ in $\mathbb{R}^m$.

**Proof:** Note that $w \in N(A) \iff Aw = 0$. Therefore if $r$ is a row vector of $A$, then $r \cdot w = 0$. This proves that $N(A) \perp \bar{\mathcal{R}}(A)$.
For the next one, do the same with $A^T$. ∎

From the Projection theorem, above lemma and Rank-Nullity Theorem, we have the following orthogonal decomposition.

**Theorem:** For any $m \times n$ matrix $A$, we have

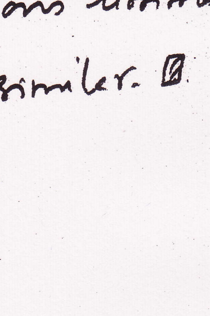
(1). $N(A) \oplus \bar{\mathcal{R}}(A) = \mathbb{R}^n$ & (2). $N(A^T) \oplus C(A) = \mathbb{R}^m$.

**Proof:** Since $\dim(\bar{\mathcal{R}}(A)) = \dim(\text{row space}) = \dim(\text{column space}) = \dim(\text{Range of } A)$, $\dim(\bar{\mathcal{R}}(A)) + \dim N(A) = n$ by Rank-Nullity Theorem. Therefore, $\dim N(A) = \dim \bar{\mathcal{R}}(A)^{\perp}$. Since $N(A) \subseteq \bar{\mathcal{R}}(A)^{\perp}$ by previous lemma, we have $N(A) = \bar{\mathcal{R}}(A)^{\perp}$. The proof of the other one is similar. ∎

Sometimes the above result is called FT of LA.

**An Application:** (Least square solutions)

**Problem:** Consider the system $Ax = b$. If $b \in C(A)$ = range of $A$, then the system is consistent and has a solution. If $b \notin C(A)$ or the system is inconsistent, then we find a "pseudo" solution or best possible solution in some sense.




**Least square solution:** Let $A = (a_{ij})_{m\times n} : \mathbb{R}^n \to \mathbb{R}^m$.

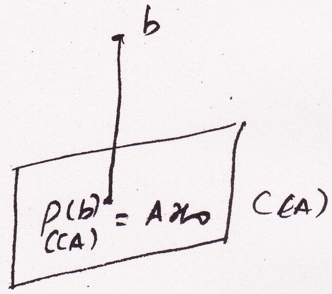
Find $x_0 \in \mathbb{R}^n$ s.t. $\|Ax_0 - b\| = \min_{x \in \mathbb{R}^n} \|Ax - b\|$.

We will show that such $x_0$ exists (need not be unique, though $Ax_0$ is unique). The element $x_0$ is called a least square solution. We will use the FT of LA to find $x_0$ explicitly.

<u>What is $x_0$:</u> Since $C(A)$ or the range of $A$ is a subspace of $\mathbb{R}^m$, $A(x_0)$ has to be the projection of $b$ over $C(A)$ ie $P_{C(A)}(b)$. Since $P_{C(A)}(b) \in C(A)$, $\exists\, x_0 \in \mathbb{R}^n$ s.t. $Ax_0 = P_{C(A)}(b)$. So $x_0$ exists. Note that $x_0 + N(A)$ - is the least square solution set.

<u>How to find $Ax_0$ & $x_0$:</u> one way is to find the orthogonal projection of $b$ over $C(A)$. To find the projection, one has to find an orthonormal basis of $C(A)$ using G-S process. Such a computation may be cumbersome. Here we will take a bypass using the FT of LA to find $x_0$

<u>The method:</u> Note that by FT of LA,




$$\mathbb{R}^n = C(A) \oplus N(A^T) \Rightarrow b = P_{C(A)}(b) - \left(P_{C(A)}(b) - b\right)$$
$$= Ax_0 - (Ax_0 - b)$$
$$\Rightarrow A^T(Ax_0 - b) = 0 \quad \text{as } Ax_0 - b \in N(A^T)$$
$$\Rightarrow A^TAx_0 = A^Tb$$

$\Rightarrow$ $x_0$ is a solution of the normal equation $\boxed{A^TAx_0 = A^Tb}$

<u>Remark:</u> (1) We have shown that if $x_0$ is a least square solution then it satisfies the equation: $A^TAx = A^Tb$. One can go back in the previous argument & show the converse. So we have:
$x_0$ is a least square solution $\Leftrightarrow$ $x_0$ satisfies $A^TAx = A^Tb$.

(2) Note that $Ax - b$ may not be zero for some $x$; but $A^T(Ax-b) = 0$ for some $x_0$. So, instead of solving the original equation $Ax = b$, we solve the equation: $A^TAx = A^Tb$.