Hilbert Projection Theorem

Let H be a Hilbert space (inner product space that is complete with respect to the norm induced by the inner product) and M be a finite dimensioal subspace of H. Then for any $x \in H$, there exists a unique $y \in M$ such that

$$\min_{m \in M} \|x - m\|$$

has a unique solution y. In other words "we can find a unique point in M that is closest to x". If m^* is the closest point to x in M, then $x - m^* \perp M$.

Proof: See lecture notes.

> **Remark:** The proof stated that $m^* = x_1$ is the closest point to M. It can also be interpreted as the best approximation of x choosen from the vectors in M. The x_2 term is the error in the approximation.

Example: Let $V = \mathbb{R}^2$ and $M = \mathrm{span}\{[1,1]^T\}$. Find the best approximation of $x = [4,7]^T$ in M.

Solution: We need to find m^* such that $m^* = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\|x - m^*\|$ is minimum.

$$(x - m^*) \perp M \implies < x - m^*, m >= 0 \quad \forall m \in M$$

$$< x - m^*, \begin{bmatrix} 1 \\ 1 \end{bmatrix} >= 0$$

$$\text{Replace } x \text{ and } m^* \text{ with their values.}$$

$$< \begin{bmatrix} 4 - \alpha \\ 7 - \alpha \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} >= 0$$

$$\text{Recall that } < x, y >= x^T y.$$

$$4 - \alpha + 7 - \alpha = 0$$

$$\alpha = \frac{11}{2}$$

$$m^* = \frac{11}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Example: Let $x \in V$ and $M = \mathrm{span}\{v_1, v_2\}$. Find the best approximation of x in M.

Solution: We need to find m^* such that $m^* = \alpha_1 v_1 + \alpha_2 v_2$ that is in the span of M and $\|x - m^*\|$ is minimum.

$$(x - m^*) \perp M \implies (x - m^*) \perp \text{both } v_1 \text{ and } v_2$$

$$< x - \alpha_1 v_1 - \alpha_2 v_2, v_1 >=< x, v_1 > -\alpha_1 < v_1, v_1 > -\alpha_2 < v_2, v_1 >= 0$$

$$< x - \alpha_1 v_1 - \alpha_2 v_2, v_2 >=< x, v_2 > -\alpha_1 < v_1, v_2 > -\alpha_2 < v_2, v_2 >= 0$$

$$\begin{bmatrix} < v_1, v_1 > & < v_2, v_1 > \\ < v_1, v_2 > & < v_2, v_2 > \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} < x, v_1 > \\ < x, v_2 > \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} < v_1, v_1 > & < v_2, v_1 > \\ < v_1, v_2 > & < v_2, v_2 > \end{bmatrix}^{-1} \begin{bmatrix} < x, v_1 > \\ < x, v_2 > \end{bmatrix}$$

$$m^* = \alpha_1 v_1 + \alpha_2 v_2$$

Example: Let $V = \mathbb{R}^3$ and $M = \mathrm{span}\{[1,1,1]^T, [1,0,1]^T\}$. Find the best approximation of $x = [4,7,2]^T$ in M.

Solution: We need to find m^* such that $m^* = \alpha_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ that is in the span of M and $\|x - m^*\|$ is minimum.

$$(x - m^*) \perp M \implies (x - m^*) \perp \text{both } v_1 \text{ and } v_2$$

$$\begin{bmatrix} < v_1, v_1 > & < v_2, v_1 > \\ < v_1, v_2 > & < v_2, v_2 > \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} < x, v_1 > \\ < x, v_2 > \end{bmatrix}$$

$$\text{Recall that } < x, y >= x^T y.$$

$$\begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 13 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 13 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 3/2 \end{bmatrix} \begin{bmatrix} 13 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 7 \\ -4 \end{bmatrix}$$

$$m^* = \alpha_1 v_1 + \alpha_2 v_2 = 7 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 4 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 3 \end{bmatrix}$$

<u>Example:</u> Let H be the space of square integrable functions on $[-\pi, \pi]$ with inner product $< f, g >= \int_{-\pi}^{\pi} f(t)\bar{g(t)} dt$. Let M be the subspace of H, $M = \text{span}\{e^{jkt}/\sqrt{2\pi}\}$, k from $-N$ to N. Note that dimension of M is $2N+1$.

$$< f_n, f_m >= \int_{-\pi}^{\pi} \frac{e^{jnt}}{\sqrt{2\pi}} \frac{e^{-jmt}}{\sqrt{2\pi}} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(n-m)t} dt$$

$$\text{If } n \neq m, \text{ then } \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(n-m)t} dt = \frac{1}{2\pi} \frac{e^{j(n-m)t}}{j(n-m)} \Big|_{-\pi}^{\pi} = 0$$

$$\text{If } n = m, \text{ then } \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(n-m)t} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1 dt = 1$$

$$\text{Therefore, } < f_n, f_m >= \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

Now, let $g \in H$ be an arbitrary function. Then $g = g_1 + g_2$ where $g_1 \in M$ and $g_2 \in M^\perp$. We need to find g_1 such that $\|g - g_1\|$ is minimum.

$$g_1 = \sum_{k=-N}^{N} \alpha_k \frac{e^{jkt}}{\sqrt{2\pi}} \text{ where } \alpha_k =< g, f_k >= \int_{-\pi}^{\pi} g(t) \frac{e^{-jkt}}{\sqrt{2\pi}} dt$$

$$\alpha_k = \int_{-\pi}^{\pi} g(t) \frac{e^{-jkt}}{\sqrt{2\pi}} dt$$

$$g_1 = \sum_{k=-N}^{N} \int_{-\pi}^{\pi} g(t) \frac{e^{-jkt}}{\sqrt{2\pi}} \frac{e^{jkt}}{\sqrt{2\pi}} dt = \sum_{k=-N}^{N} \int_{-\pi}^{\pi} g(t) \frac{1}{2\pi} dt = \int_{-\pi}^{\pi} g(t) dt$$

<u>Example:</u> Find the orthagonal projection of the vector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ onto the subspace $M = \text{span}\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \}$.

<u>Solution:</u> We know that the projection matrix is $P = B(B^T B)^{-1} B^T$ where B is the matrix whose columns are the basis vectors of M. Therefore,

$$B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$B^T B = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3/8 & -1/8 \\ -1/8 & 3/8 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

$$P \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \end{bmatrix}$$

Solution of Linear Equations

Consider the linear equation expressed as

$$Ax = b \text{ where } A \in \mathbb{C}^{m \times n}, x \in \mathbb{C}^n, b \in \mathbb{C}^m$$

If $m = n$, then the equation has a unique solution. If $m < n$, then the equation has infinitely many solutions. If $m > n$, then the equation has no solution.

> - A solution exists if and only if $b \in \text{range}(A)$.
> - A solution is unique if and only if A is full rank, or equivalently, A has linearly independent columns, or $N(A) = \{0\}$.

Example: Let $A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $b = \begin{bmatrix} 2.2 \\ 1.9 \\ 2.1 \\ 1.8 \end{bmatrix}$. Find x such that $Ax = b$.

Solution:

$$b \in^? \text{range}(A)$$

$$\text{range}(A) = \text{span}\{[1, 1, 1, 1]^T\}$$

$$\text{span}\{[1, 1, 1, 1]^T\} = \{[a, a, a, a]^T \mid a \in \mathbb{R}\}$$

$$b \notin \text{range}(A)$$

Therefore, there is no exact solution. We need to find the best approximation of b in $\text{range}(A)$.

Lets call this best approximation b^*. Then, $b^* = \alpha \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $||Ax - b^*||^2$ is minimum.

$$B = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [4]^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

$$Pb = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [4]^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2.2 \\ 1.9 \\ 2.1 \\ 1.8 \end{bmatrix}$$

$$Pb = b^* = \begin{bmatrix} 2.0 \\ 2.0 \\ 2.0 \\ 2.0 \end{bmatrix}$$

Example: Let $A = \begin{bmatrix} 2 & 1 \\ 2 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{bmatrix}$. Find x such that $Ax = b$.

Solution: Now, the rows of A are linearly dependent. Therefore, A is not full column rank. Therefore, there is no exact solution. We need to find the best approximation of b in $\mathrm{range}(A)$.

$$b^* = \begin{bmatrix} \bar{l} \\ \bar{l} \\ \bar{l} \\ \bar{l} \end{bmatrix} \text{ where } \bar{l} = \frac{l_1 + l_2 + l_3 + l_4}{4}$$

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \bar{l} \\ \bar{l} \\ \bar{l} \\ \bar{l} \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix} \text{ is any solution, more examples } \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \ldots$$

$$\text{A minimum norm solution is can be found by minimizing the norm of } x.$$

$$\text{For that, we can project any solution onto the null space perpendicular to } A.$$

$$x_1 = \mathrm{Proj}_{N(A)^\perp} \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix}$$

$$\text{Corollary: } N(A)^\perp = \mathrm{range}(A^*)$$

$$A^* = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\mathrm{range}(A^*) = \mathrm{span}\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

$$\mathrm{Proj}_{N(A)^\perp} \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix} = \mathrm{Proj}_{\mathrm{range}(A^*)} \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix}$$

$$\mathrm{Proj}_{\mathrm{range}(A^*)} \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix} = \frac{< \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} >}{< \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} >} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathrm{Proj}_{\mathrm{range}(A^*)} \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix} = \frac{\bar{l}}{5} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\text{Or in projection matrix form, } Q = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ then, } P = Q(Q^*Q)^{-1}Q^*$$

$$P = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 2 & 1 \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$x_1 = Px = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix} \begin{bmatrix} \bar{l} \\ -\bar{l} \end{bmatrix} = \begin{bmatrix} 4/5\bar{l} - 2/5\bar{l} \\ 2/5\bar{l} - 1/5\bar{l} \end{bmatrix} = \begin{bmatrix} 2/5\bar{l} \\ 1/5\bar{l} \end{bmatrix} = \bar{l}/5 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Special Cases of $Ax = b$

1.1 Columns of A are linearly independent

If the columns of A are linearly independent, then A is full rank and $A^T A$ is invertible. Therefore, $x = (A^T A)^{-1} A^T b$ is the unique solution.

A is full column rank with A is $m \times n$ and $m \geq n$ if and only if $N(A) = \{0\}$. (Tall matrix)

If $b \in \mathrm{range}(A)$, then the solution exists and is unique. If $b \notin \mathrm{R}(A)$, then the solution does not exist.

$Ax = \mathrm{Proj}_{\mathrm{R}(A)} b$ is the best approximation of b in $\mathrm{R}(A)$.

$P = A(A^T A)^{-1} A^T$ is the projection matrix onto $\mathrm{R}(A)$.

$Ax = A(A^T A)^{-1} A^T b$ is the best approximation of b in $\mathrm{R}(A)$.

$Ax - A(A^T A)^{-1} A^T b = 0$ since the projection of b onto $\mathrm{R}(A)$.

$A(x - (A^T A)^{-1} A^T b) = 0$

$x - (A^T A)^{-1} A^T b \in N(A)$ and the null space contains only the zero vector.

$x = (A^T A)^{-1} A^T b$ is the unique solution.

1.2 Columns of A are linearly dependent

If the columns of A are linearly dependent, then A is not full rank and $A^T A$ is not invertible. Therefore, $x = (A^T A)^{-1} A^T b$ is not the unique solution.

2.1 Rows of A are linearly independent

If the rows of A are linearly independent, then A is full row rank and AA^T is invertible. Therefore, $x = A^T (AA^T)^{-1} b$ is the unique solution.

A is full row rank with A is $m \times n$ and $m \leq n$ if and only if $N(A^T) = \{0\}$. (Wide matrix)

If $b \in \mathrm{range}(A)$, then the solution exists and is unique. If $b \notin \mathrm{R}(A)$, then the solution does not exist.

$\dim(\mathrm{R}(A)) = \dim(\mathrm{R}(A^T)) = \mathrm{rank}(A) = \mathrm{rank}(A^T) \implies b \in \mathrm{R}(A)$

For a minimum norm solution, we need to project b onto the null space perpendicular to A.

$x = \mathrm{Proj}_{N(A)^\perp} b = \mathrm{Proj}_{\mathrm{R}(A^*)} b$

$P = A^*(AA^*)^{-1} A = A^*(A^* A)^{-1} A$

$Px = A^*(AA^*)^{-1} Ax = A^*(AA^*)^{-1} b$

2.2 Rows of A are linearly dependent

If the rows of A are linearly dependent, then A is not full row rank and AA^T is not invertible. Therefore, $x = A^T (AA^T)^{-1} b$ is not the unique solution.

3.1 A is square and invertible

If A is square and invertible, then $x = A^{-1} b$ is the unique solution.

Example: Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. Find x such that $Ax = b$.

Solution:

First, we need to check if $b \in \mathrm{range}(A)$. Since A has linearly dependent columns

$$\mathrm{range}(A) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} \right\}$$

$$\dim(\mathrm{range}(A)) = \mathrm{rank}(A) = 2$$

$$\dim(N(A)) = n - \text{rank}(A) = 3 - 2 = 1$$

$$N(A) = \text{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \right\}$$

$$b \notin \text{range}(A)$$

Therefore, there is no exact solution. We need to find the best approximation of b in $\text{range}(A)$.

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix}$$

$$P = B(B^T B)^{-1} B^T$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 11 & 7 \\ 7 & 11 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \end{bmatrix}$$

$$P = \frac{1}{72} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 11 & -7 \\ -7 & 11 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \end{bmatrix}$$

$$P = \frac{1}{72} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & -10 & 26 \\ 4 & 26 & -10 \end{bmatrix}$$

$$P = \frac{1}{72} \begin{bmatrix} 8 & 16 & 16 \\ 16 & 68 & -4 \\ 16 & -4 & 68 \end{bmatrix}$$

$$Pb = \frac{1}{72} \begin{bmatrix} 8 & 16 & 16 \\ 16 & 68 & -4 \\ 16 & -4 & 68 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 13/6 \\ -5/6 \end{bmatrix}$$

$$b^* = Pb = \begin{bmatrix} 1/3 \\ 13/6 \\ -5/6 \end{bmatrix}$$

Now, the problem has a solution. However for the uniqueness, we need to check if A is full row rank. Since A has linearly dependent rows, A is not full row rank. Therefore, the solution is not unique, hence we need to find the minimum norm solution. For that, we need to project b onto the null space perpendicular to A.

Starting with any solution x,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\text{Let } x_1 = 0$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 13/6 \\ -5/6 \end{bmatrix}$$

$$x_2 = \frac{-7}{6} \text{ and } x_3 = \frac{9}{6}$$

$$x = \begin{bmatrix} 0 \\ -7/6 \\ 9/6 \end{bmatrix}$$

Now, x_{min} is the projection of x onto the null space perpendicular to A.

$$x_{min} = \text{Proj}_{N(A)^{\perp}} x = \text{Proj}_{\text{R}(A^*)} x$$

$$A^* = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 1 & 3 & 1 \end{bmatrix}$$

$$\mathrm{range}(A^*) = \mathrm{span}\left\{ \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix} \right\}$$

$$Q = \begin{bmatrix}1 & 1\\1 & 2\\1 & 3\end{bmatrix}$$

$$P = Q(Q^*Q)^{-1}Q^*$$

$$P = \begin{bmatrix}1 & 1\\1 & 2\\1 & 3\end{bmatrix} \begin{bmatrix}3 & 6\\6 & 14\end{bmatrix}^{-1} \begin{bmatrix}1 & 1 & 1\\1 & 2 & 3\end{bmatrix}$$

$$P = \frac{1}{6}\begin{bmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{bmatrix}$$

$$x_{min} = Px = \frac{1}{6}\begin{bmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{bmatrix} \begin{bmatrix}0\\-7/6\\9/6\end{bmatrix} = \begin{bmatrix}-25/36\\4/36\\31/36\end{bmatrix}$$

#EE501 - `Linear Systems Theory` at `METU`