# 3.3 Projections and Least Squares

A system of equations $Ax = b$ has a solution iff $b \in C(A)$.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1}$$

Equation (1) can be written as

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

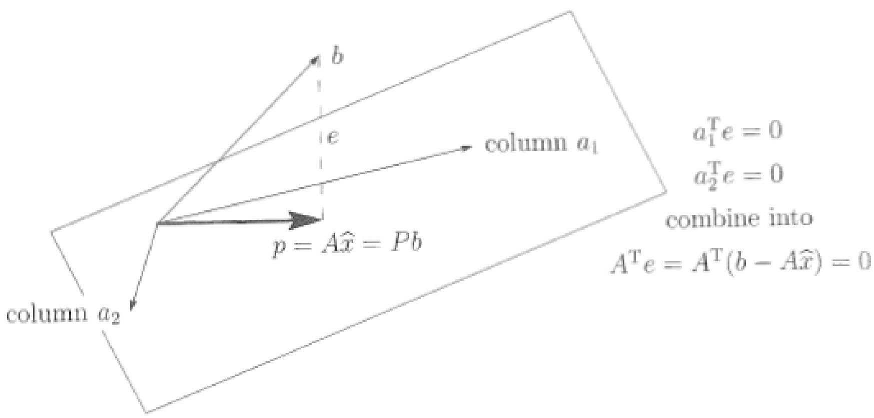
Let the co-efficient vector of $x_i$ be $a_i$. Now, $b \in C(A)$,i.e, there exists $c_1, c_2, \cdots, c_n$ such that $b = \sum_{i=1}^{n} c_i a_i$. Then for $x_i = c_i,\ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is a solution of (1).

If $x$ is a solution of (1), then $c_i = x_i$, i.e., $b$ is a linear combination of columns of $A$, i.e., $b \in C(A)$.

Consider a system of linear equations

$$Ax = b \tag{2}$$

with $b \notin C(A)$. Then (2) has no solution. Then one must choose $\widehat{x}$ a best fit solution of it.



As $b \notin C(A)$, let us obtain a column vector $p$ in $C(A)$, which is the most closest vector of $b$ in $C(A)$ such that the error $e = b - p$ is of minimum length .i.e,$\| e \|^2 = \| b - p \|^2$ is minimum. Hence it is called least square approximation.

$$\begin{aligned}
&\| e \|^2 \quad \text{is minimum} \\
&\Longleftrightarrow e \perp C(A) \\
&\Longleftrightarrow e \in N(A^T) \\
&\Longleftrightarrow A^T e = 0 \\
&\Longleftrightarrow A^T(b - p) = 0 \\
&\Longleftrightarrow A^T(b - A\widehat{x}) = 0 \\
&\Longleftrightarrow A^T b = A^T A\widehat{x}
\end{aligned}$$

i.e., Normal equations $A^T A\widehat{x} = A^T b$.
Best estimate: $\widehat{x} = (A^T A)^{-1} A^T b$.
Projection $p = A\widehat{x} = A(A^T A)^{-1} A^T b$.
Projection matrix $P = A(A^T A)^{-1} A^T$.

**Remark 1** *1. Suppose $b$ is actually in the column space of $A$ it is a combination $b = Ax$ of the columns. Then the projection of $b$ is still $b$: $b$ in column space*

$$p = A(A^T A)^{-1} A^T Ax = Ax = b:$$

*The closest point $p$ is just $b$ itself, which is obvious.*

*2. At the other extreme, suppose $b$ is perpendicular to every column, so $A^T b = 0$. In this case $b$ projects to the zero vector: $b$ in left nullspace, i.e., $b \in N(A^T)$*

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

*3. When $A$ is square and invertible, the column space is the whole space. Every vector projects to itself, $p$ equals $b$, and $\widehat{x} = x$ : If $A$ is invertible*

$$p = A(A^T A)^{-1} A^T b = AA^{-1}(A^T)^{-1} A^T b = b:$$

*This is the only case when we can take apart $(A^T A)^{-1}$, and write it as $A^{-1}(A^T)^{-1}$. When $A$ is rectangular that is not possible.*

*4. Suppose $A$ has only one column, containing $a$. Then the matrix $A^T A$ is the number $a^T a$ and $\widehat{x}$ is $a^T b / a^T a$.*

5. *$A^TA$ has the same nullspace as $A$.*

   *If $Ax = 0$, then $A^TAx = 0$. Vectors $x$ in the nullspace of $A$ are also in the nullspace of $A^TA$. To go in the other direction, start by supposing that $A^TAx = 0$, and take the inner product with $x$ to show that $Ax = 0$ :*

   $$x^TA^TAx = 0 \Rightarrow \| Ax \|^2 = 0 \Rightarrow Ax = 0.$$

   *Hence, the two nullspaces are identical.*

6. *If $A$ has independent columns , then $N(A^TA) = N(A) = \{0\}$.*

7. *If $A$ has independent columns, then $A^TA$ is square, symmetric, and invertible.*

# Projection Matrices

$$\begin{aligned} P &= \text{Projection matrix, which projects a vector onto C(A)} \\ &= A(A^TA)^{-1}A^T \end{aligned}$$

i.e, $p = Pb$.
Note that

$$\begin{aligned} P^2 &= A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T \\ &= A(A^TA)^{-1}(A^TA)(A^TA)^{-1}A^T \\ &= A(A^TA)^{-1}A^T \\ &= P \end{aligned}$$

and

$$\begin{aligned} P^T &= [A(A^TA)^{-1}A^T]^T \\ &= (A^T)^T[(A^TA)^{-1}]^TA^T \\ &= A(A^TA)^{-1}A^T \\ &= P. \end{aligned}$$

Suppose that $A$ is invertible. Then,

$$\begin{aligned} P &= A(A^TA)^{-1}A^T \\ &= AA^{-1}(A^T)^{-1}A^T \\ &= I. \end{aligned}$$

Hence, $p = Pb = Ib = b$. Therefore, $error = p - b = 0$.

Exercise-3.3

1. Solve $Ax = b$ by least squares and find $p = A\widehat{x}$, if $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

   Verify that the error $b - p$ is perpendicular to the columns of $A$.

   **Solution:**

$$\begin{aligned} \widehat{x} &= (A^T A)^{-1} A^T b \\ &= \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} P &= A\widehat{x} \\ &= (A^T A)^{-1} A^T b \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}. \end{aligned}$$

Here, Error$= b - p = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{-2}{3} \end{pmatrix}$.

$b - p$ is perpendicular to the column of the matrix.

$$(b-p)^T C_1 = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{-2}{3} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{2}{3} - \frac{2}{3} = 0 \text{ and}$$

$$(b-p)^T C_2 = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{-2}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \frac{2}{3} - \frac{2}{3} = 0.$$

4. The following system has no solution: $Ax = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 9 \end{pmatrix} = b.$

   Sketch and solve a straight-line fit that leads to the minimization of the quadratic $(C-D-4)^2 + (C-5)^2 + (C+D-9)^2$. What is the projection of b onto the column space of $A$?

   **Solution:**

$$\begin{aligned} \widehat{x} &= (A^T A)^{-1} A^T b \\ &= \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 9 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{2} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 9 \end{pmatrix} \\ &= \begin{pmatrix} 6 \\ \frac{5}{2} \end{pmatrix}. \end{aligned}$$

   Now,

$$\begin{aligned} P &= A\widehat{x} \\ &= \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ \frac{5}{2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{7}{2} \\ 6 \\ \frac{17}{2} \end{pmatrix}. \end{aligned}$$

# Assignments

Exercise-3.3, Q. 2,9,12,24.