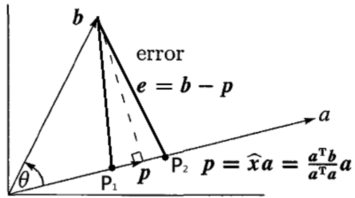
EE 205003 Session 17

Che Lin

Institute of Communications Engineering

Department of Electrical Engineering

Projection onto a line



Q: How do we find a point $\mathbf{p}$
on the line (determined by vector $\mathbf{a}$)
that is closest to $\mathbf{b}$?

$\mathbf{p}$: intersection of a line through $\mathbf{b}$
that is orthogonal to $\mathbf{a}$
($\mathbf{P}_1$, $\mathbf{P}_2$ have longer distance)

More precisely

Think of $\mathbf{p}$ as an approximation of $\mathbf{b}$,
then $\mathbf{e} = \mathbf{b} - \mathbf{p}$ is the error vector

Since $\mathbf{p}$ is along the line of $\mathbf{a}$

$\Rightarrow \mathbf{p} = \widehat{x}\mathbf{a}$ for some $\widehat{x}$

Also, $\mathbf{a} \perp \mathbf{e}$

$$\Rightarrow \mathbf{a}^\intercal(\mathbf{b} - \mathbf{p}) = \mathbf{a}^\intercal(\mathbf{b} - \widehat{x}\mathbf{a}) = 0$$

$$\Rightarrow \mathbf{a}^\intercal\mathbf{a}\widehat{x} = \mathbf{a}^\intercal\mathbf{b} \Rightarrow \widehat{x} = \frac{\mathbf{a}^\intercal\mathbf{b}}{\mathbf{a}^\intercal\mathbf{a}}$$

Now, we have

$$\mathbf{p} = \widehat{x}\mathbf{a} = \mathbf{a}\widehat{x} = \mathbf{a}\frac{\mathbf{a}^\intercal\mathbf{b}}{\mathbf{a}^\intercal\mathbf{a}}$$

(doubling $\mathbf{b}$ doubles $\mathbf{p}$, doubling $\mathbf{a}$ dose NOT affect $\mathbf{p}$)

Projection Matrix $(\mathbf{p} = P\mathbf{b})$

$$\mathbf{p} = \mathbf{a}\frac{\mathbf{a}^\intercal \mathbf{b}}{\mathbf{a}^\intercal \mathbf{a}} = \underbrace{\frac{\mathbf{a}\mathbf{a}^\intercal}{\mathbf{a}^\intercal \mathbf{a}}}_{P}\mathbf{b}$$

(For 3D)

$\mathbf{a}$ 3×3

$$P = \frac{\mathbf{a}\mathbf{a}^\intercal}{\mathbf{a}^\intercal \mathbf{a}}$$

$\mathbf{a}\mathbf{a}^\intercal$ ← rank-one matrix

$\mathbf{a}^\intercal \mathbf{a}$ ← a number

(Procedure : Find $\widehat{x} \rightarrow \mathbf{p} \rightarrow P$)

Special case I : If $\mathbf{b} = \mathbf{a}$, $\widehat{x} = 1$

$\Rightarrow P\mathbf{a} = \mathbf{a}$ (projection of $\mathbf{a}$ onto $\mathbf{a}$ is itself)

Special case II : If $\mathbf{b} \perp \mathbf{a}$, $\mathbf{a}^\intercal \mathbf{b} = 0$

$\Rightarrow \mathbf{p} = \mathbf{0}$

Projection Matrix ($\mathbf{p} = P\mathbf{b}$) (cont.)

Note 1: col. space of P is spanned by $\mathbf{a}$

($\because$ for any $\mathbf{b}$, $P\mathbf{b}$ lies on the line determined by $\mathbf{a}$)

Note 2: rank$(P) = 1$

Note 3 : P is symmetric

$(P^\intercal = (\frac{\mathbf{a}\mathbf{a}^\intercal}{\mathbf{a}^\intercal\mathbf{a}})^\intercal = \frac{1}{\mathbf{a}^\intercal\mathbf{a}}(\mathbf{a}\mathbf{a}^\intercal)^\intercal = \frac{\mathbf{a}\mathbf{a}^\intercal}{\mathbf{a}^\intercal\mathbf{a}} = P)$

Note 4 : $P^2 = P$

($P^2\mathbf{b} = P\mathbf{b}$ or $P(P\mathbf{b}) = P\mathbf{b}$

$\because$ projection of a vector already on $\mathbf{a}$ is itself)

Note 5 : $I - P$ is also a projection

($(I - P)\mathbf{b} = \mathbf{b} - \mathbf{p} = \mathbf{e}$ in the left nullspace of $\mathbf{a}$ $\because$ $\mathbf{a}^\intercal\mathbf{e} = 0$)

(P : project onto one subspace

$I - P$: project onto the perpendicular subspace)

Q : Why project ?

$A\mathbf{x} = \mathbf{b}$ may have no solution

always in col. space of A

unlikely that $\mathbf{b} \in \mathbf{C}(A)$

If not, project $\mathbf{b}$ onto $\mathbf{p} \in \mathbf{C}(A)$
then solve $A\widehat{\mathbf{x}} = \mathbf{p}$

Projection onto a Subspace

Projection onto a plane (in $\mathbb{R}^3$)

If $\mathbf{a}_1, \mathbf{a}_2$ are basis of a plane
$\Rightarrow$ the plane is $\mathbf{C}(A)$ of $A = \left[\mathbf{a}_1, \mathbf{a}_2\right]$

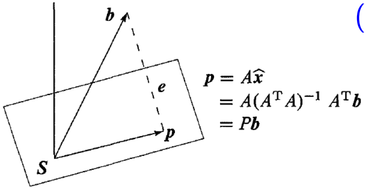
In general, for a subspace $\mathbf{S} \subseteq \mathbb{R}^m$ with
n independent basis $\mathbf{a}_1, \cdots, \mathbf{a}_n$
$\Rightarrow$ subspace is $\mathbf{C}(A)$ of $A_{m \times n} = \left[\mathbf{a}_1, \cdots, \mathbf{a}_n\right]$

Ch 4.2 Projections

Problem : Find $\mathbf{p}$ in $\mathbf{S}$ closest to $\mathbf{b}$

Since $\mathbf{p} \in \mathbf{C}(A)$, $\mathbf{p} = A\widehat{\mathbf{x}} = \widehat{x}_1\mathbf{a}_1 + \cdots + \widehat{x}_n\mathbf{a}_n$

(want to find $\widehat{x}_i$)



$\mathbf{p}$ closest to $\mathbf{b}$

$\Rightarrow \mathbf{e} = \mathbf{b} - \mathbf{p} \perp \mathbf{S}$

or $\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}} \perp \mathbf{S}$

$\Rightarrow \mathbf{e} = \mathbf{b} - A\widehat{\mathbf{x}}$ perpendicular with $\mathbf{a}_1, \cdots, \mathbf{a}_n$

$$\Rightarrow \begin{array}{c} \mathbf{a}_1^\intercal(\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \\ \mathbf{a}_2^\intercal(\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \\ \vdots \\ \mathbf{a}_n^\intercal(\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \end{array} \quad \text{or} \quad \begin{bmatrix} -\mathbf{a}_1^\intercal- \\ \vdots \\ -\mathbf{a}_n^\intercal- \end{bmatrix} \begin{bmatrix} \mathbf{b} - A\widehat{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \end{bmatrix}$$

$$\Rightarrow A^\intercal(\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \Rightarrow A^\intercal A\widehat{\mathbf{x}} = A^\intercal\mathbf{b}$$

$\|$

$\mathbf{e}$ (in the null space of A)

Find $\widehat{\mathbf{x}}$

$$\widehat{\mathbf{x}} = (A^{\intercal}A)^{-1}A^{\intercal}\mathbf{b}$$

(Q: Is $A^{\intercal}A$ invertible ?
Yes, if n columns of A are linear independent)
(will prove this later)

Find $\mathbf{p}$

$$\mathbf{p} = A\widehat{\mathbf{x}} = \underbrace{A(A^{\intercal}A)^{-1}A^{\intercal}}_{\text{projection matrix } P = A(A^{\intercal}A)^{-1}A^{\intercal}}\mathbf{b}$$

(Find $\widehat{\mathbf{x}} \rightarrow \mathbf{p} \rightarrow P$)

Alternative derivation

1. our subspace is $\mathbf{C}(A)$
2. error vector $\mathbf{e} = \mathbf{b} - A\widehat{\mathbf{x}} \perp \mathbf{C}(A)$
3. so $\mathbf{e}$ in left null space of A
   ($\mathbf{C}(A)$ and $\mathbf{N}(A^\intercal)$) are orthogonal components)
   $\Rightarrow A^\intercal\mathbf{e} = A^\intercal(\mathbf{b} - A\widehat{\mathbf{x}}) = \mathbf{0}$

($\mathbf{b}$ splitted into $\mathbf{p}$ & $\mathbf{e}$)
$(\in \mathbf{C}(A))$ $(\in \mathbf{N}(A^\intercal))$

Special cases

1. $\mathbf{b} \perp \mathbf{C}(A)$: $\mathbf{b} \in \mathbf{N}(A^\intercal)$ & $P\mathbf{b} = \mathbf{0}$
2. $\mathbf{b} \in \mathbf{C}(A)$: $A\mathbf{x} = \mathbf{b}$ for some $\mathbf{x}$ & and $P\mathbf{b} = \mathbf{b}$

Q: Can we further simplify $P = A(A^\intercal A)^{-1}A^\intercal$?

$$\begin{aligned} P = A(A^\intercal A)^{-1}A^\intercal &= A(A^{-1}(A^\intercal)^{-1})A^\intercal \\ &= (AA^{-1})((A^\intercal)^{-1}A^\intercal) \\ &= I \text{ ?} \end{aligned}$$

Wrong ! A is rectangular $\Rightarrow$ A has no inverse matrix !

$\boxed{\text{Fact}}$ $P = P^\intercal$, $P^2 = P$ (still true for general $\mathbf{v}$)

distance from $\mathbf{b}$ to subspace $= \|\mathbf{e}\|$

$\boxed{\text{Fact}}$ A is invertible iff A has linear independent columns

pf : First, we want to show that $A^\intercal A$ and A have same nullspace

if $\mathbf{x}$ is in $\mathbf{N}(A)$, then $A\mathbf{x} = \mathbf{0}$

$\Rightarrow A^\intercal A\mathbf{x} = A^\intercal(\mathbf{0}) = \mathbf{0}$

$\Rightarrow \mathbf{x}$ in $\mathbf{N}(A^\intercal A)$

if $\mathbf{x}$ in $\mathbf{N}(A^\intercal A)$, then $A^\intercal A\mathbf{x} = \mathbf{0}$

$\Rightarrow \mathbf{x}^\intercal A^\intercal A\mathbf{x} = \mathbf{x}^\intercal\mathbf{0} = 0$

$\Rightarrow \|A\mathbf{x}\|^2 = 0 \Rightarrow A\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} \in \mathbf{N}(A)$

So A & and $A^\intercal A$ have same nullspace

Now, if A has independent columns

then rank$(A) = n \Rightarrow \mathbf{N}(A) = \{\mathbf{0}\}$

$\Rightarrow \mathbf{N}(A^\intercal A) = \{\mathbf{0}\} \Rightarrow A^\intercal A$ is invertible

If $A^\intercal A$ is invertible, then $A^\intercal A$ has independent columns

$\Rightarrow \mathbf{N}(A^\intercal A) = \{\mathbf{0}\} \Rightarrow \mathbf{N}(A) = \{\mathbf{0}\}$

$\Rightarrow A$ has independent columns

Ex3 : (on p.211, textbook)

If $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$, find $\widehat{\mathbf{x}}$, $\mathbf{p}$ & P

Normal equation :

$A^\intercal A \widehat{\mathbf{x}} = A^\intercal \mathbf{b}$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \widehat{\mathbf{x}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \Rightarrow \widehat{\mathbf{x}} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

Ex3 : (on p.211, textbook) (cont.)

$$\mathbf{p} = A\widehat{\mathbf{x}} = 5\begin{bmatrix}1\\1\\1\end{bmatrix} - 3\begin{bmatrix}0\\1\\2\end{bmatrix} = \begin{bmatrix}5\\2\\-1\end{bmatrix}$$

$$\mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix}1\\-2\\1\end{bmatrix}$$ (indeed $\perp$ both columns of A)

To find $\mathbf{p}$ for every $\mathbf{b}$, we need P

$P = A(A^\intercal A)^{-1}A^\intercal$

$$(A^\intercal A)^{-1} = \tfrac{1}{6}\begin{bmatrix}5 & -3\\-3 & 3\end{bmatrix} \Rightarrow P = \begin{bmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{bmatrix}$$

(check : $P\mathbf{b} = \mathbf{p}$ & $P^2 = P$, $P^\intercal = P$)