Orthogonal Projection

Hung-yi Lee

Reference

- Textbook: Chapter 7.3, 7.4

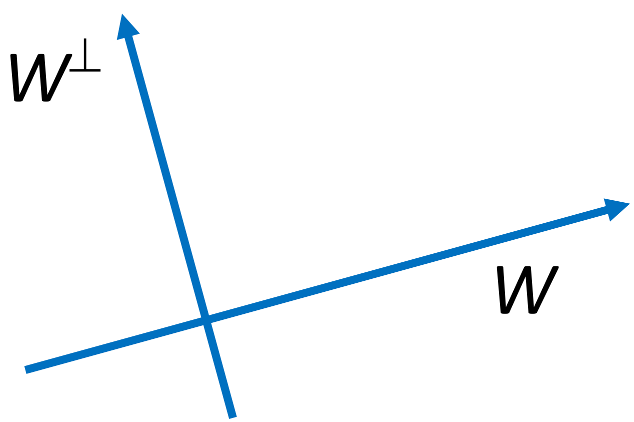
Orthogonal Projection

- What is Orthogonal Complement
- What is Orthogonal Projection
- How to do Orthogonal Projection
- Application of Orthogonal Projection

Orthogonal Complement

- The orthogonal complement of a nonempty vector set $\mathcal{S}$ is denoted as $\mathcal{S}^{\perp}$ (S perp).
- $\mathcal{S}^{\perp}$ is the set of vectors that are orthogonal to every vector in $\mathcal{S}$

$$S^{\perp} = \{v: v \cdot u = 0, \forall u \in S\}$$



$\mathcal{S} = \mathcal{R}^n$

$\mathcal{S} = \{0\}$

Orthogonal Complement

- The orthogonal complement of a nonempty vector set $\mathcal{S}$ is denoted as $\mathcal{S}^{\perp}$ (S perp).
- $\mathcal{S}^{\perp}$ is the set of vectors that are orthogonal to every vector in $\mathcal{S}$

$$S^{\perp} = \{v : v \cdot u = 0, \forall u \in S\}$$

$$W = \left\{ \begin{bmatrix} w_1 \\ w_2 \\ 0 \end{bmatrix} \middle| w_1, w_2 \in \mathcal{R} \right\}$$

$$V = \left\{ \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix} \middle| v_3 \in \mathcal{R} \right\} = W^{\perp}?$$

$V \subseteq W^{\perp}$:

for all $\mathbf{v} \in V$ and $\mathbf{w} \in W$, $\mathbf{v} \bullet \mathbf{w} = 0$

$W^{\perp} \subseteq V$:

since $\mathbf{e}_1, \mathbf{e}_2 \in W$, all $\mathbf{z} = [\ z_1\ \ z_2\ \ z_3\ ]^T \in W^{\perp}$ must have $z_1 = z_2 = 0$

Properties of Orthogonal Complement

Is $S^{\perp}$ always a subspace?

For any nonempty vector set S, $(Span\ S)^{\perp} = S^{\perp}$

Let W be a subspace, and B be a basis of W.

$B^{\perp} = W^{\perp}$

What is $S \cap S^{\perp}$?

Zero vector

Properties of Orthogonal Complement

- Example:

For W = Span{$\mathbf{u}_1$, $\mathbf{u}_2$}, where $\mathbf{u}_1 = [\ 1\ \ 1\ \ -1\ \ 4\ ]^T$ and $\mathbf{u}_2 = [\ 1\ -1\ \ 1\ \ 2\ ]^T$

$\mathbf{v} \in W^{\perp}$ if and only if $\mathbf{u}_1 \bullet \mathbf{v} = \mathbf{u}_2 \bullet \mathbf{v} = \mathbf{0}$

i.e., $\mathbf{v} = [\ x_1\ \ x_2\ \ x_3\ \ x_4\ ]^T$ satisfies

$$\begin{aligned} x_1 + x_2 - x_3 + 4x_4 &= 0 \\ x_1 - x_2 + x_3 + 2x_4 &= 0. \end{aligned} \quad \Leftrightarrow \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3x_4 \\ x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

$\Leftrightarrow \mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $W^{\perp}$.

$$A = \begin{bmatrix} 1 & 1 & -1 & 4 \\ 1 & -1 & 1 & 2 \end{bmatrix}$$

$W^{\perp}$ = Solutions of "Ax=0" = Null A

Properties of Orthogonal Complement

- For any matrix A

$$(Row\ A)^{\perp} = Null\ A$$

$\mathbf{v} \in (\text{Row } A)^{\perp}$

$\Leftrightarrow A\mathbf{v} = \mathbf{0}.$

$$(Col\ A)^{\perp} = Null\ A^{T}$$

$(\text{Col } A)^{\perp} = (\text{Row } A^{T})^{\perp} = \text{Null } A^{T}.$

For any subspace W of R^n

$$\underbrace{dim W}_{\text{rank}} + \underbrace{dim W^{\perp}}_{\text{nullity}} = n$$

Unique

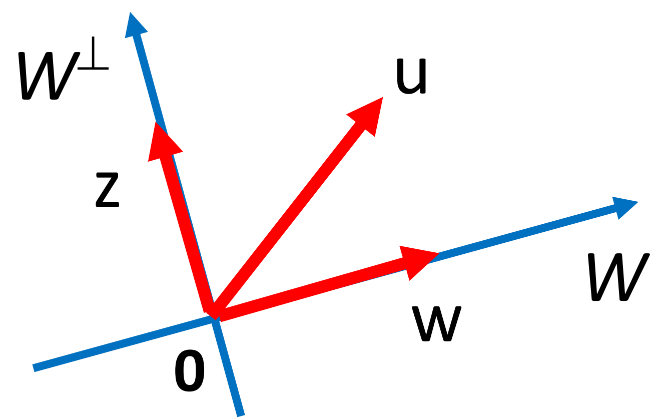
For any subspace W of R^n $\quad dimW + dimW^{\perp} = n$

Basis: $\{w_1, w_2, \cdots, w_k\}$ Basis: $\{z_1, z_2, \cdots, z_{n-k}\}$

Basis for R^n

For ***every*** vector u,

$\underbrace{w}_{\in W}$ and $\underbrace{z}_{\in W^{\perp}}$: $u = w + z$ ***(unique)***



Orthogonal Projection

- What is Orthogonal Complement
- What is Orthogonal Projection
- How to do Orthogonal Projection
- Application of Orthogonal Projection

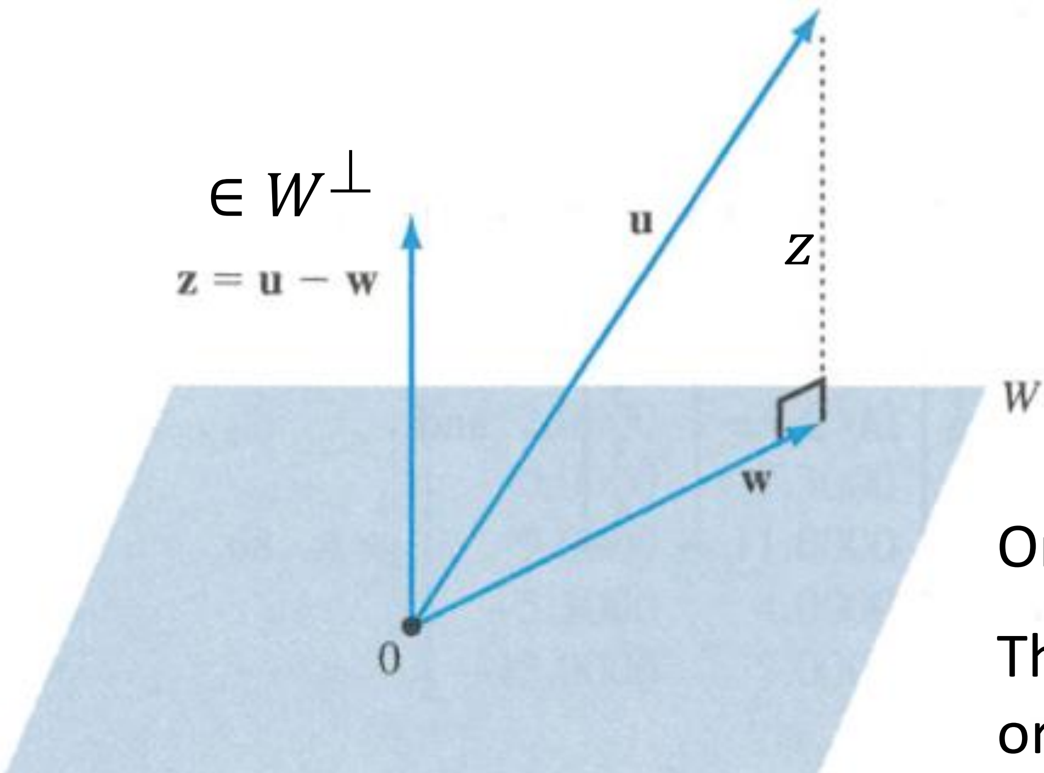
Orthogonal Projection

orthogonal projection

u = w + z ***(unique)***

$\in W$ $\in W^{\perp}$

$\in W^{\perp}$

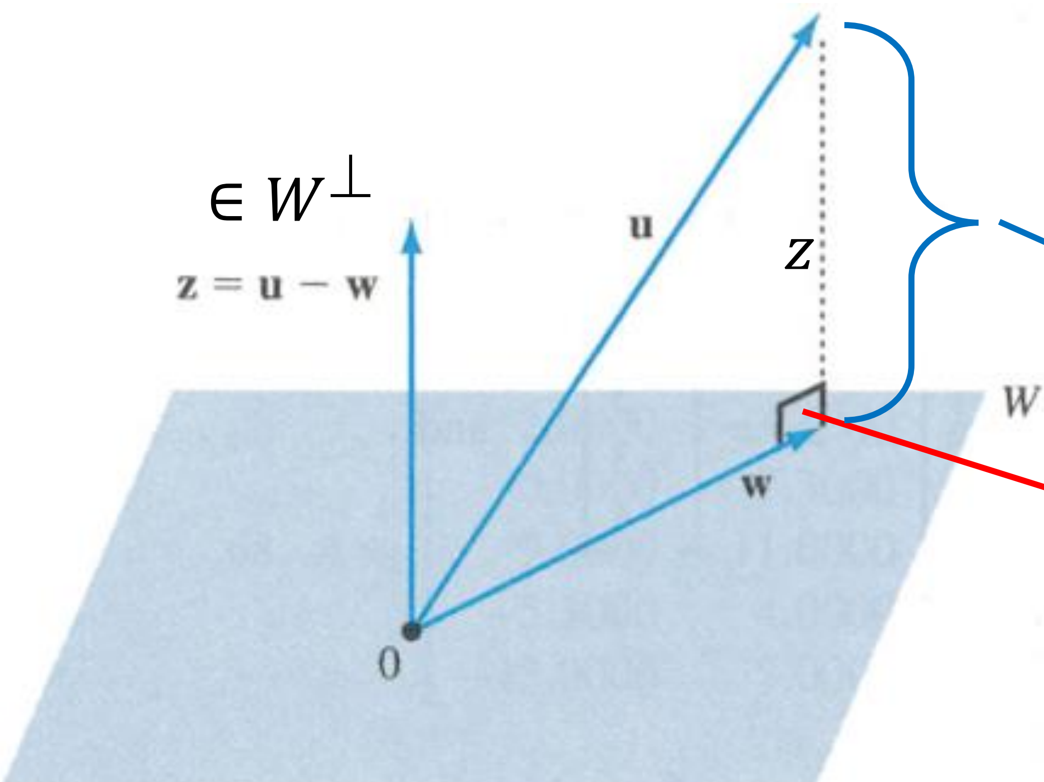


Orthogonal Projection Operator:

The function $U_W(u)$ is the orthogonal projection of u on W.

Linear?

Orthogonal Projection



w in subspace *W* is "closest" to vector **u**.

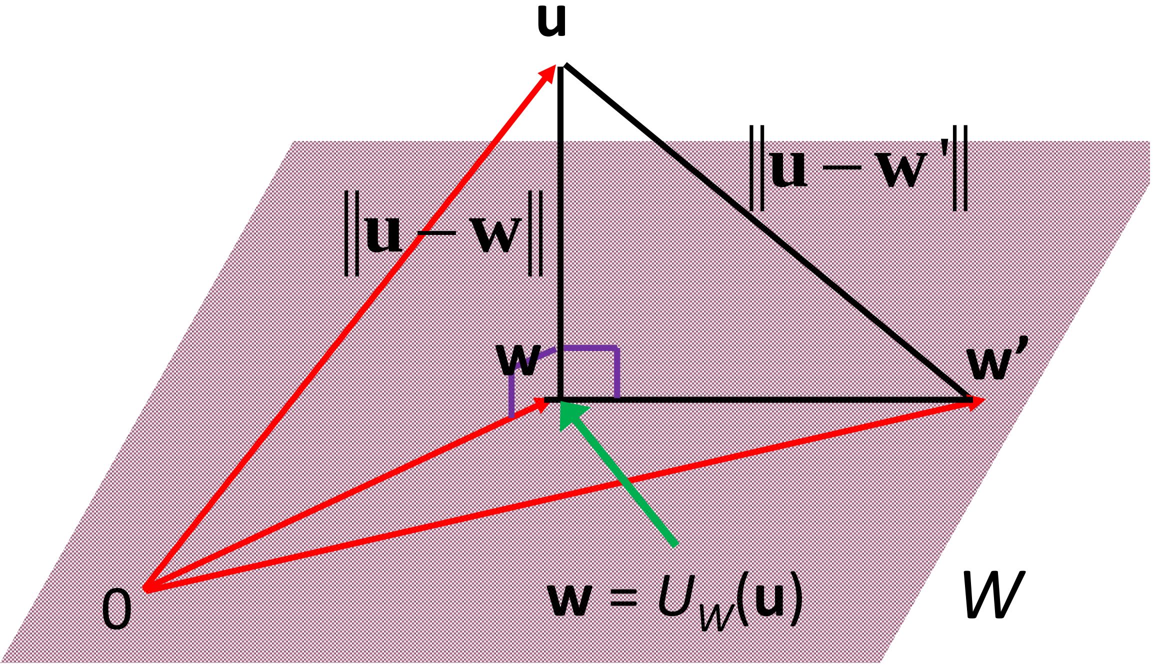
"closest"

Always orthogonal

z is always orthogonal to all vectors in *W*.

Closest Vector Property

- Among all vectors in subspace W, the vector closest to u is the orthogonal projection of u on W



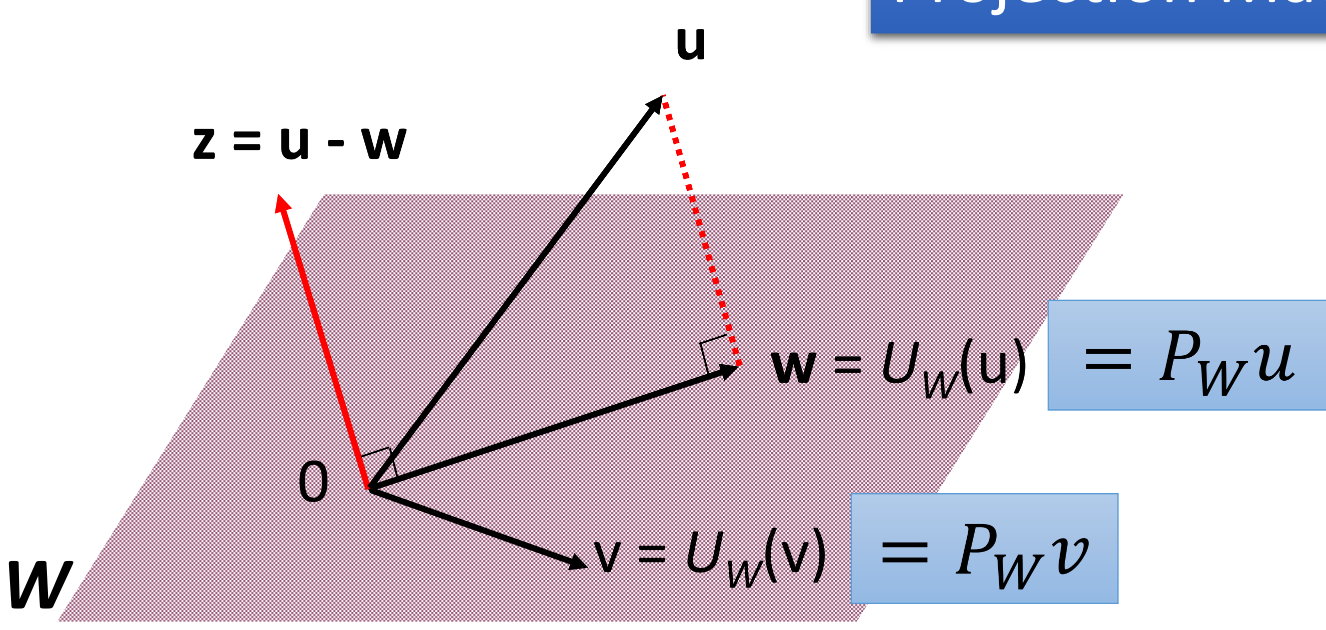
$\forall\ \mathbf{w}, \mathbf{w}', \mathbf{w} - \mathbf{w}' \in W.$

The distance from a vector u to a subspace W is the distance between u and the orthogonal projection of u on W

Orthogonal Projection Matrix

Orthogonal projection operator is linear.

➡ It has standard matrix. ➡ Orthogonal Projection Matrix P_w



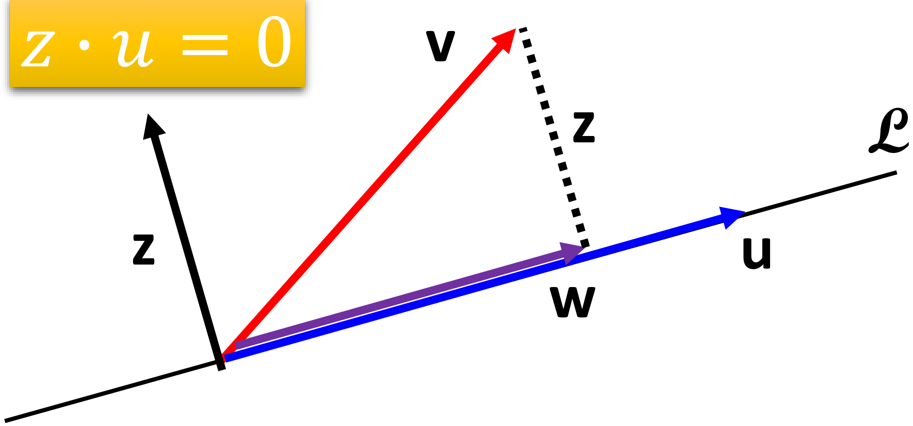
Orthogonal Projection

- What is Orthogonal Complement
- What is Orthogonal Projection
- How to do Orthogonal Projection
- Application of Orthogonal Projection

Orthogonal Projection on a line

- Orthogonal projection of a vector on a line

$z \cdot u = 0$



v: any vector

u: any nonzero vector on $\mathcal{L}$

w: orthogonal projection of **v** onto $\mathcal{L}$, **w** = c**u**

z: **v** – **w**

$$(v - w) \cdot u \;= (v - cu) \cdot u \;= v \cdot u - cu \cdot u \;= v \cdot u - c\|u\|^2$$

=0

$$c = \frac{v \cdot u}{\|u\|^2} \qquad w = cu = \frac{v \cdot u}{\|u\|^2} u$$

Distance from tip of **v** to $\mathcal{L}$: $\|z\| = \|v - w\| \;= \left\| v - \frac{v \cdot u}{\|u\|^2} u \right\|$

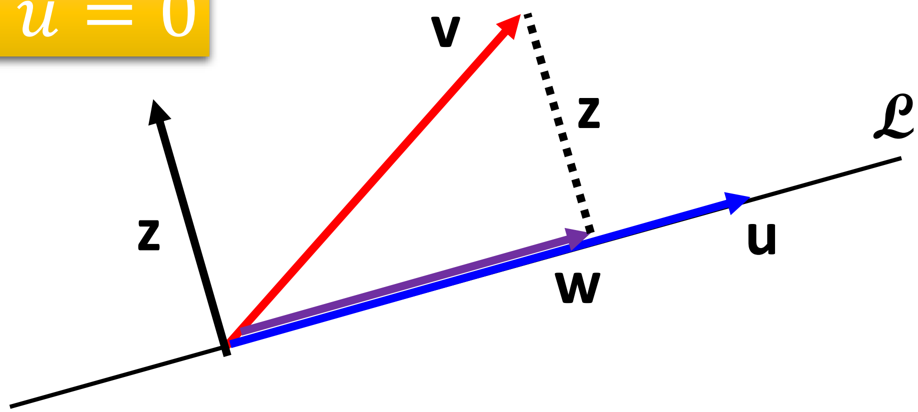
Orthogonal Projection

$$c = \frac{v \cdot u}{\|u\|^2}$$

$$w = cu = \frac{v \cdot u}{\|u\|^2} u$$

- Example:

$\boxed{z \cdot u = 0}$



$\mathcal{L}$ is $y = (1/2)x$

$$v = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Orthogonal Projection Matrix

- Let C be an n x k matrix whose columns form a basis for a subspace W

$$P_W = C(C^T C)^{-1} C^T$$

n x n

Proof: Let $\mathbf{u} \in \mathcal{R}^n$ and $\mathbf{w} = U_W(\mathbf{u})$.

Orthogonal Projection Matrix

- Let C be an n x k matrix whose columns form a basis for a subspace W

$$P_W = C(C^TC)^{-1}C^T$$

n x n

Let C be a matrix with linearly independent columns. Then C^TC is invertible.

Orthogonal Projection Matrix

- Example: Let W be the 2-dimensional subspace of $\mathcal{R}^3$ with equation $x_1 - x_2 + 2x_3 = 0$.

$$P_W = C(C^T C)^{-1} C^T$$

W has a basis $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ $\quad C = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

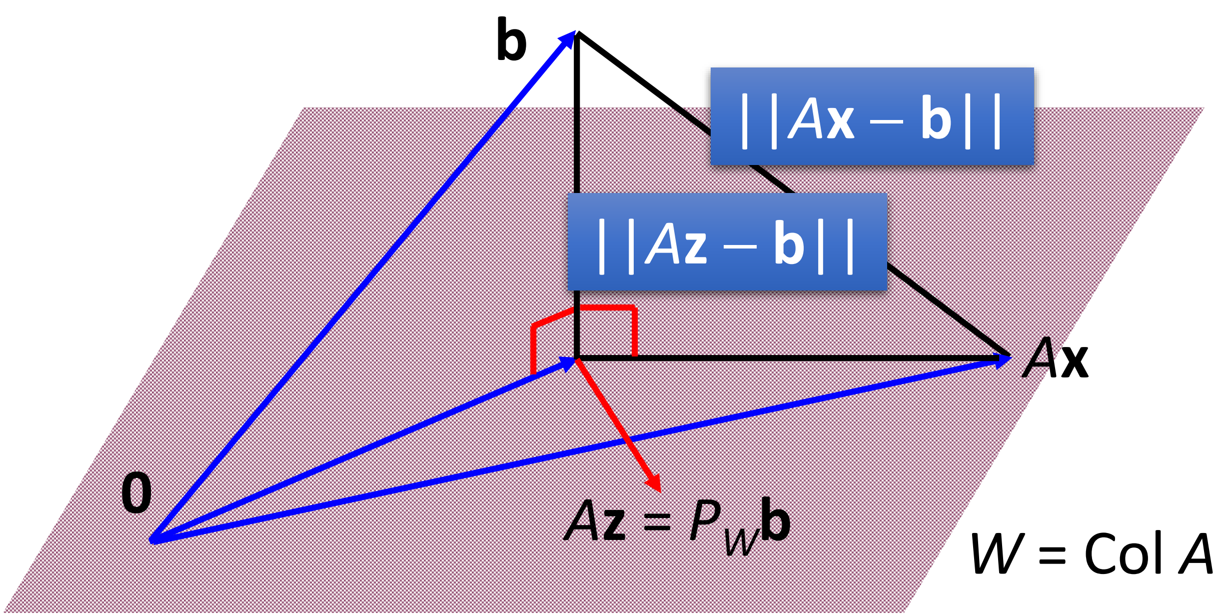
$$P_W = \frac{1}{6} \begin{bmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix} \qquad P_W \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}$$

Orthogonal Projection

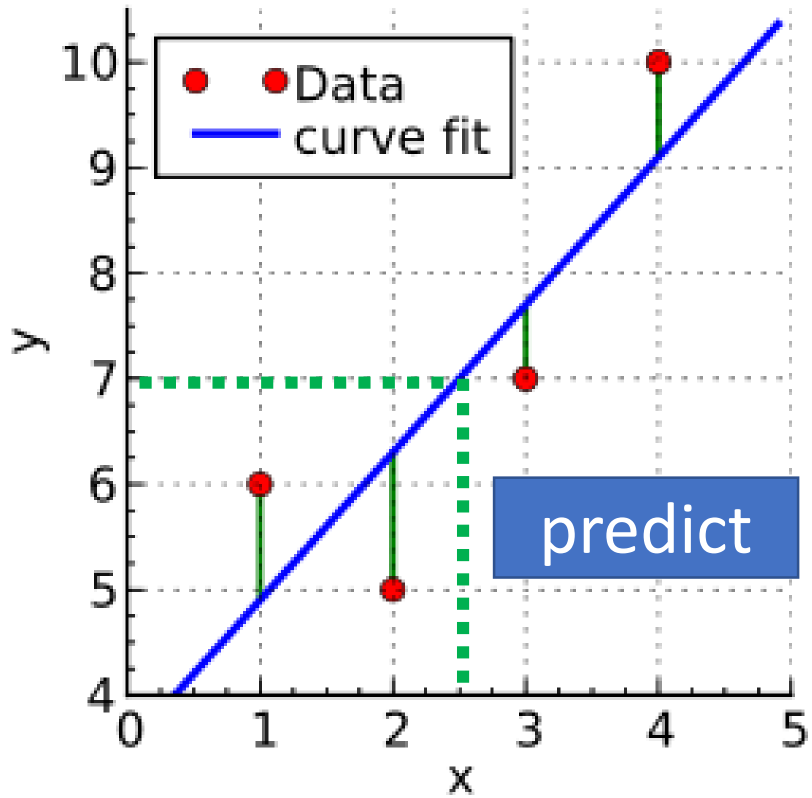
- What is Orthogonal Complement
- What is Orthogonal Projection
- How to do Orthogonal Projection
- Application of Orthogonal Projection

Solution of Inconsistent System of Linear Equations

- Suppose $A\mathbf{x} = \mathbf{b}$ is an inconsistent system of linear equations.
- $\mathbf{b}$ is not in the column space of A
- Find vector $\mathbf{z}$ minimizing $||A\mathbf{z} - \mathbf{b}||$



Least Square Approximation



data pairs:

$x_1 \rightarrow y_1$

$x_2 \rightarrow y_2$

$\vdots$

$x_i \rightarrow y_i$

$\vdots$

e.g.

(今天股票,明天股票)

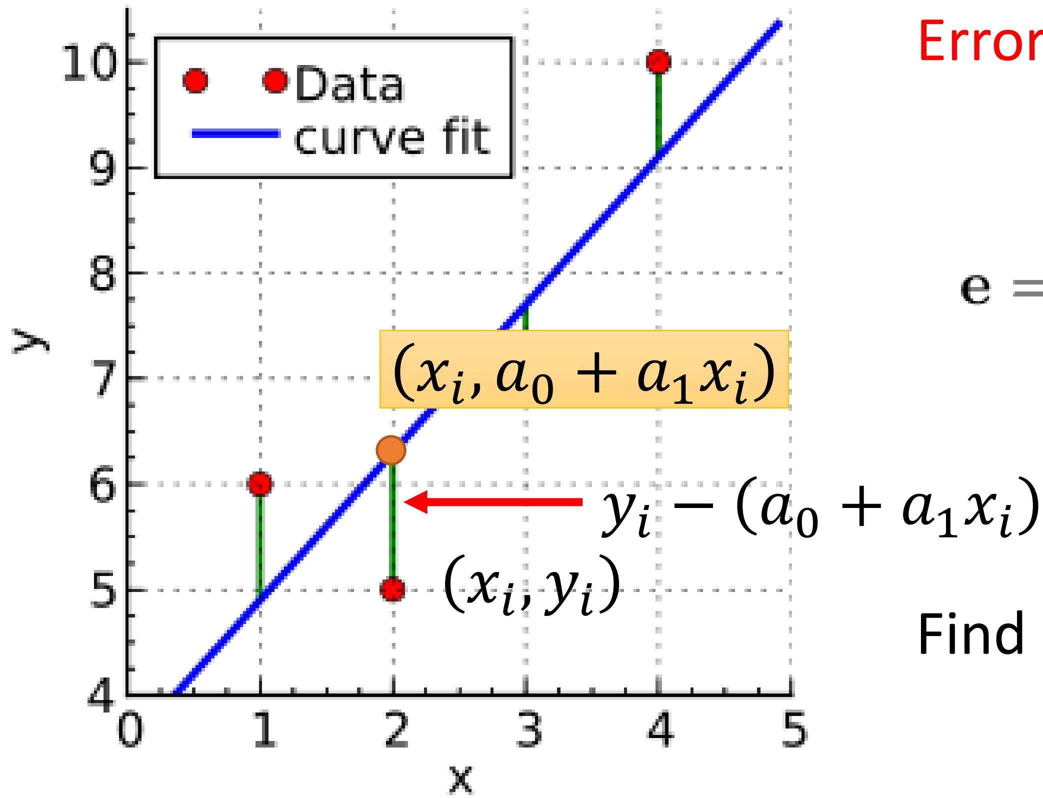
(今天PM2.5,明天PM2.5)

Find the "least-square line" $y = a_0 + a_1 x$ to best fit the data

Regression

Least Square Approximation

$y = a_0 + a_1 x$



Error Vector:

$$\mathbf{e} = \begin{bmatrix} y_1 - (a_0 + a_1 x_1) \\ y_2 - (a_0 + a_1 x_2) \\ \vdots \\ y_n - (a_0 + a_1 x_n) \end{bmatrix}$$

Find a_0 and a_1 minimizing E

$$E = ||\mathbf{e}||^2$$

$$E = [y_1 - (a_0 + a_1 x_1)]^2 + [y_2 - (a_0 + a_1 x_2)]^2 + \cdots + [y_n - (a_0 + a_1 x_n)]^2$$

Least Square Approximation

Error Vector:

$$\mathbf{e} = \begin{bmatrix} \boxed{y_1} \; \boxed{- (a_0} + \boxed{a_1 x_1)} \\ \boxed{y_2} \; \boxed{- (a_0} + \boxed{a_1 x_2)} \\ \vdots \\ \boxed{y_n} \; \boxed{- (a_0} + \boxed{a_1 x_n)} \end{bmatrix}$$

$$\mathbf{e} = \underline{\mathbf{y}} - \underline{a_0 \mathbf{v}_1} - \underline{a_1 \mathbf{v}_2}$$

$$\boxed{\mathbf{y} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}, \boxed{\mathbf{v}_1 \triangleq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}, \boxed{\mathbf{v}_2 \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}$$

Find a_0 and a_1 minimizing E

$$E = ||\mathbf{e}||^2$$

$$C \triangleq \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}, \text{ and } \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

$$E = ||\mathbf{y} - (a_0\mathbf{v}_1 + a_1\mathbf{v}_2)||^2 = ||\mathbf{y} - C\mathbf{a}||^2$$

Least Square Approximation

Find **a** minimizing

$$E = \|\mathbf{y} - C\mathbf{a}\|^2$$

$\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ (L.I.)

$$\mathbf{y} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{v}_1 \triangleq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \mathbf{v}_2 \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Ca is the orthogonal projection of **y** on W = Span $\mathscr{B}$.

find **a** such that $C\mathbf{a} = P_W\mathbf{y}$

$$C \triangleq \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}, \text{ and } \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = (C^T C)^{-1} C^T \mathbf{y}$$

Example 1

Rough weight x_i (in pounds)	Finished weight y_i (in pounds)
2.60	2.00
2.72	2.10
2.75	2.10
2.67	2.03
2.68	2.04

$$C = \begin{bmatrix} 1 & 2.60 \\ 1 & 2.72 \\ 1 & 2.75 \\ 1 & 2.67 \\ 1 & 2.68 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 2.00 \\ 2.10 \\ 2.10 \\ 2.03 \\ 2.04 \end{bmatrix}$$

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = (C^T C)^{-1} C^T \mathbf{y} \approx \begin{bmatrix} 0.056 \\ 0.745 \end{bmatrix}$$

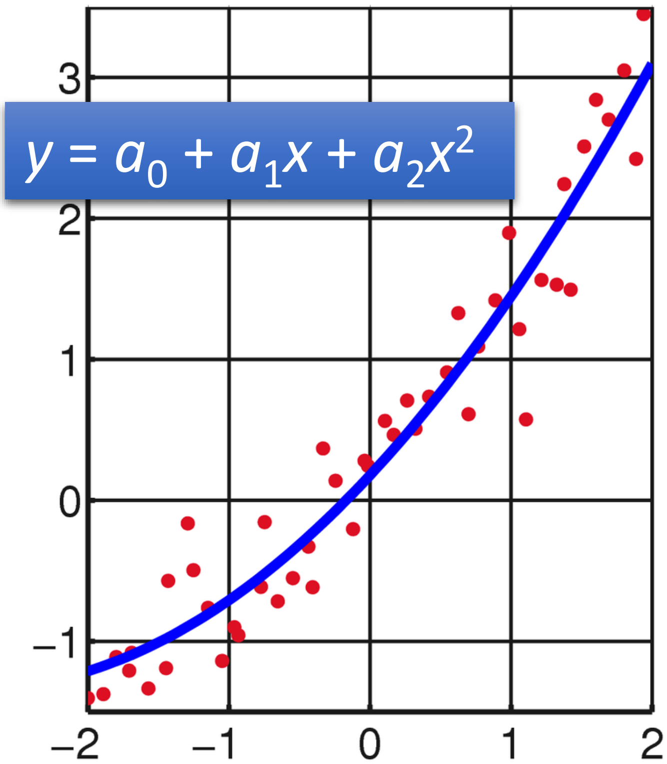
$\Rightarrow y = 0.056 + 0.745x.$

Prediction:
if the rough weight is 2.65,
the finished weight is
0.056 +0.745(2.65) = 2.030.

(estimation)

Least Square Approximation

- Best quadratic fit: using $y = a_0 + a_1 x + a_2 x^2$ to fit the data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$



$$e = \begin{bmatrix} y_1 - (a_0 + a_1 x_1 + a_2 x_1^2) \\ y_2 - (a_0 + a_1 x_2 + a_2 x_2^2) \\ \vdots \\ y_n - (a_0 + a_1 x_n + a_2 x_n^2) \end{bmatrix}$$

Find a_0, a_1 and a_2 minimizing E

$$E = ||\mathbf{e}||^2$$

Least Square Approximation

- Best quadratic fit: using $y = a_0 + a_1x + a_2x^2$ to fit the data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^n \end{bmatrix}$$

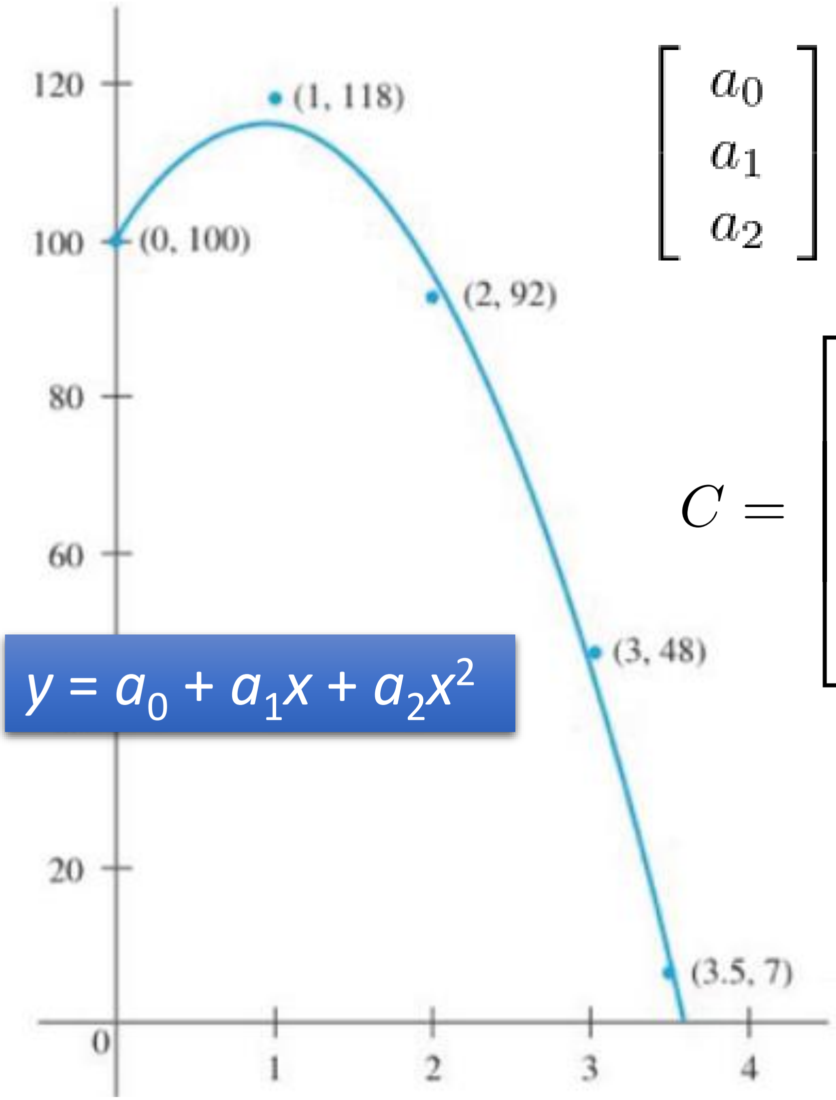
$$e = \begin{bmatrix} y_1 - (a_0 + a_1x_1 + a_2x_1^2) \\ y_2 - (a_0 + a_1x_2 + a_2x_2^2) \\ \vdots \\ y_n - (a_0 + a_1x_n + a_2x_n^2) \end{bmatrix}$$

$$C = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$$

Find a_0, a_1 and a_2 minimizing E

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = (C^TC)^{-1}C^T\mathbf{y}.$$

$$E = ||\mathbf{e}||^2$$



$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = (C^TC)^{-1}C^T\mathbf{y}$$

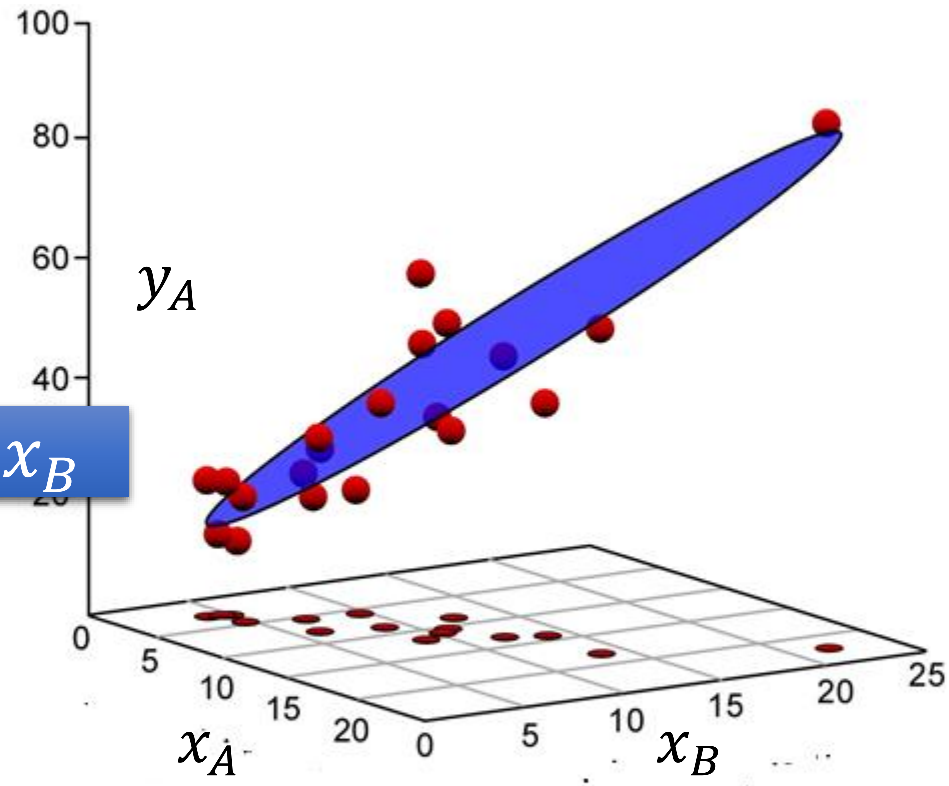
$$C = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 3.5 & 12.25 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 100 \\ 118 \\ 92 \\ 48 \\ 7 \end{bmatrix}$$

$$y = 101.00 + 29.77x - 16.11x^2$$

Best fitting polynomial of any desired maximum degree may be found with the same method.

Multivariable Least Square Approximation

$$y_A = a_0 + a_1 x_A + a_2 x_B$$



http://www.palass.org/publications/newsletter/palaeomath-101/palaeomath-part-4-regression-iv