6-4 : The Projection Matrix

basis

Given $W = sp(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k) \subset \mathbb{R}^n$, given $\vec{b} \in \mathbb{R}^n$

$$\Rightarrow \underset{\exists!}{\vec{b}} = \vec{b}_W + \vec{b}_{W^\perp} = \underbrace{\vec{p}}_{\in W} + \underbrace{(\vec{b} - \vec{p})}_{\in W^\perp}$$

If T: linear transformation, s.t. $T(\vec{b}) = \vec{b}_W$. (ch 6-2, get $\vec{b}_W$ by G-S of $\{\vec{a}_i\}$) dot product

Q: What's the s.m.r. of T?

$\triangle$

① $\because \vec{b}_W \in W \quad \therefore \exists\, r_1, r_2, \dots, r_k \in \mathbb{R}$ s.t. $\vec{b}_W = r_1\vec{a}_1 + r_2\vec{a}_2 + \dots + r_k\vec{a}_k = A\vec{r}$

where $A_{n\times k} = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \\ | & | & & | \end{bmatrix}$, $\vec{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}$

② $\vec{b}_{W^\perp} = \vec{b} - \vec{b}_W = \vec{b} - A\vec{r} \in W^\perp$

③ $\forall \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{R}^k$, $A\vec{x} \in W \quad \therefore \underbrace{(A\vec{x})}_{W} \circ \underbrace{(\vec{b} - A\vec{r})}_{W^\perp} = 0$

$$0 = (A\vec{x}) \circ (\vec{b} - A\vec{r}) = \vec{x}^T A^T (\vec{b} - A\vec{r}) = \vec{x}^T (A^T\vec{b} - A^TA\vec{r}) \Rightarrow A^T\vec{b} - A^TA\vec{r} = \vec{0}$$

$\therefore A^TA\vec{r} = A^T\vec{b}$ ☆ if A^TA : invertible $\Rightarrow \vec{r} = (A^TA)^{-1}A^T\vec{b} \Rightarrow \boxed{\vec{b}_W = A\vec{r} = A(A^TA)^{-1}A^T\vec{b}}$

$\therefore T(\vec{b}) = \vec{b}_W = \underline{A(A^TA)^{-1}A^T}\vec{b} \quad \therefore$ the s.m.r. of $T = A(A^TA)^{-1}A^T = P$: projection matrix

Thm 6.10

$A_{m\times n}$, $\text{rank}(A)=r \Rightarrow \text{rank}(A^TA)=r$

p.f.

Recall

1. $(A^T_{n\times m}\ A_{m\times n})_{n\times n}$

2. $\text{rank}(A)=r \Leftrightarrow \dim(\text{null}(A))=n-r$

$\text{rank}(A^TA)=r \Leftrightarrow \dim(\text{null}(A^TA))=n-r$

rewrite the statement : $A_{m\times n}$, $\dim(\text{null}(A))=n-r \Rightarrow \dim(\text{null}(A^TA))=n-r$

① $\forall \vec{v}\in \text{null}(A)$, i.e. $A\vec{v}=\vec{0} \Rightarrow A^TA\vec{v}=A^T\vec{0}=\vec{0} \Rightarrow \vec{v}\in\text{null}(A^TA)$

$\Rightarrow \text{null}(A)\subseteq\text{null}(A^TA)$

② $\vec{v}\in\text{null}(A^TA)$, i.e. $A^TA\vec{v}=\vec{0} \Rightarrow \vec{v}^T(A^TA\vec{v})=\vec{v}^T\vec{0}=0 \Rightarrow A\vec{v}=\vec{0}$

$\vec{v}^T(A^TA\vec{v}) = (\vec{v}^TA^T)(A\vec{v})=(A\vec{v})\cdot(A\vec{v})=\|A\vec{v}\|^2 \Rightarrow \vec{v}\in\text{null}(A)$

$\therefore \text{null}(A^TA)\subseteq\text{null}(A)$

$\therefore \text{null}(A)=\text{null}(A^TA)$ $\quad\therefore \dim(\text{null}(A^TA))=\dim(\text{null}(A))=n-r$ #

Is A^TA : invertible? Yes!

Back to previous page

basis $\quad$ ✩ $\dim(W)=k$

Given $W = sp(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k) \subset \mathbb{R}^n$, given $\vec{b} \in \mathbb{R}^n$, $A_{n\times k} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \end{bmatrix} \Rightarrow \text{rank}(A) = k$

$\therefore$ by Thm 6.10 $(A^T_{k\times n} A_{n\times k})_{k\times k}$ has rank k

$\because$ rank $(A^TA) = k$, A^TA : $k\times k$ matrix $\therefore$ A^TA : invertible

$\therefore$ The $\boxed{\text{Projection Matrix } P = A(A^TA)^{-1}A^T}$

$\therefore \forall \vec{b}$, $\boxed{\vec{b}_W = P\vec{b} = A(A^TA)^{-1}A^T\vec{b}}$

ex: $W = sp(\vec{a}) = sp\left(\begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}\right)$, $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, find $\vec{b}_W = ?$

$A = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$ $\quad \therefore A^TA = \begin{bmatrix} 2 & 4 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} = [29]$

$\therefore \vec{b}_W = A(A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \frac{1}{29} \begin{bmatrix} 2 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{29}\begin{bmatrix} 4 & 8 & 6 \\ 8 & 16 & 12 \\ 6 & 12 & 9 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{19}{29}\begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$

ex:

Find the projection matrix for the y-z plane in $\mathbb{R}^3$

sol.

$W = sp(\vec{e}_2, \vec{e}_3)$, find P

$A = [\vec{e}_2 \ \vec{e}_3] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$, $A^T A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\therefore P = A(A^T A)^{-1} A^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$\therefore \forall \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \in \mathbb{R}^3$, $P\vec{b} = \begin{bmatrix} 0 \\ b_2 \\ b_3 \end{bmatrix}$

ex:

Find the projection matrix for the plane $2x - y - 3z = 0$ in $\mathbb{R}^3$

sol.

$W =$ plane $2x - y - 3z = 0$, the normal vector $\vec{n}$ of W, $\vec{n} = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix}$

pick $\vec{a}_1, \vec{a}_2 \in W$, $\vec{a}_1 = \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ $\quad \vec{a}_1 \nparallel \vec{a}_2$

$\therefore \dim(W) = 2$, $W = sp(\vec{a}_1, \vec{a}_2)$ $\quad \therefore A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}$

$$(A^TA)^{-1} = \begin{bmatrix} 10 & 6 \\ 6 & 5 \end{bmatrix}^{-1} = \frac{1}{14}\begin{bmatrix} 5 & -6 \\ -6 & 10 \end{bmatrix}$$

$$P = \frac{1}{14}\begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 5 & -6 \\ -6 & 10 \end{bmatrix}\begin{bmatrix} 0 & 3 & -1 \\ 1 & 2 & 0 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 10 & 2 & 6 \\ 2 & 13 & -3 \\ 6 & -3 & 5 \end{bmatrix}$$

$\therefore \forall \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, $\vec{b}_W = P\vec{b} = \frac{1}{14}\begin{bmatrix} 10b_1 + 2b_2 + 6b_3 \\ 2b_1 + 13b_2 - 3b_3 \\ 6b_1 - 3b_2 + 5b_3 \end{bmatrix}$

Thm

$W \subset \mathbb{R}^n$, $\exists !\ P_{n\times n}$ s.t. $\forall \vec{b} \in \mathbb{R}^n$, $P\vec{b} = \vec{b}_W$

Moreover, $P = A(A^TA)^{-1}A^T$, where $A_{n\times k}$ has columns $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k$

and $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\}$ · basis for W

p.f. (略)

$\because\ \forall \vec{b} \in \mathbb{R}$, $\vec{b}_W$: unique , T : linear transf. $T(\vec{b}) = \vec{b}_W$

Prop.

P : projection matrix $\Rightarrow \begin{cases} P^2 = P & \text{: idempotent} \\ P^T = P & \text{: symmetric} \end{cases}$

p.f. (略)

use $P = A(A^TA)^{-1}A^T$

Thm 6.12

Every projection matrix P for $W \subset \mathbb{R}^n$ $\Rightarrow$ P : idempotent & symmetric

Conversely, every $n \times n$ idempotent & symmetric matrix is a projection matrix

i.e. the project matrix for $col(P)$

$A\vec{r}$, $\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \\ | & | & & | \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}$

$\Uparrow$
$col(A)$

$= r_1\vec{a}_1 + r_2\vec{a}_2 + \cdots + r_k\vec{a}_k$

pf.

① by Prop

② let $W := col(P)$, $\forall \vec{b} \in \mathbb{R}^n$

(i) check $P\vec{b} \in W$, $\because P\vec{b} \in col(P) = W$ ✓

(ii) check $\vec{b} - P\vec{b} \in W^{\perp}$, $\forall \vec{x} \in \mathbb{R}^n$

$$\underbrace{(\vec{b} - P\vec{b})}_{? \in W^{\perp}} \cdot \underbrace{(P\vec{x})}_{\in W} = (\vec{b} - P\vec{b})^T (P\vec{x}) = (I\vec{b} - P\vec{b})^T (P\vec{x}) = ((I-P)\vec{b})^T (P\vec{x})$$

$$= \vec{b}^T \underbrace{(I-P)^T}_{P:\text{ symmetric}} (P\vec{x}) = \vec{b}^T (I-P) P\vec{x} = \vec{b}^T \underbrace{(P-P^2)}_{P:\text{ idempotent}} \vec{x} = \vec{b}^T \underbrace{O}_{\text{zero matrix}} \vec{x} = [0]$$

$\therefore \vec{b} - P\vec{b} \in W^{\perp}$ $\therefore \vec{b}_W = P\vec{b}$ #

☺ I hate $(A^TA)^{-1}$, Hope: $A^TA = I$, i.e $(A^TA)^{-1} = I$

Q: When does $A^TA = I$?

$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \\ | & | & & | \end{bmatrix}$, $A^TA = \begin{bmatrix} - \vec{a}_1^T - \\ \vdots \\ - \vec{a}_k^T - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \\ | & | & & | \end{bmatrix} = B = [b_{ij}] \Rightarrow b_{ij} = \vec{a}_i \cdot \vec{a}_j$

$\therefore A^TA = I$ iff $b_{ij} = \vec{a}_i \cdot \vec{a}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$ iff $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$: orthonormal

iff A : orthogonal

Thm

Let $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$: orthornormal basis for $W \subset \mathbb{R}^n$

$\Rightarrow$ The projection matrix $P = AA^T$, where $A_{n \times k}$ having columns $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$

ex:

if $W = sp(\vec{a}_1, \vec{a}_2) \subset \mathbb{R}^3$, $\vec{a}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, $\vec{a}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

sol

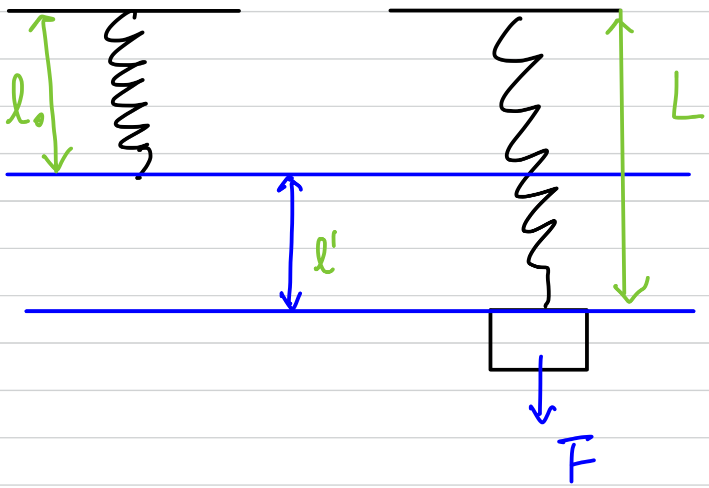
check $\vec{a}_1 \circ \vec{a}_2 = 0$, $\|\vec{a}_1\| = \|\vec{a}_2\| = 1$ $\therefore \{\vec{a}_1, \vec{a}_2\}$: orthornormal basis for W

let $A = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ -1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \end{bmatrix}$, $A^T A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

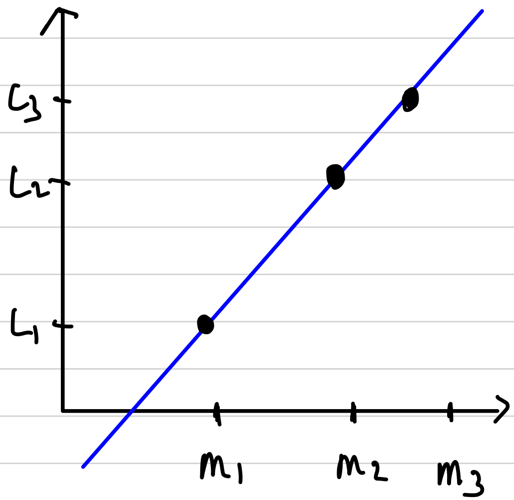
$$\therefore P = AA^T = \begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$$

§ 6.5 The Method of Least Squares

Hook's Law



$$\begin{cases} F = k\,l' , \\ L = l_0 + l' \end{cases} \Rightarrow \quad L = l_0 + \frac{1}{k} F$$



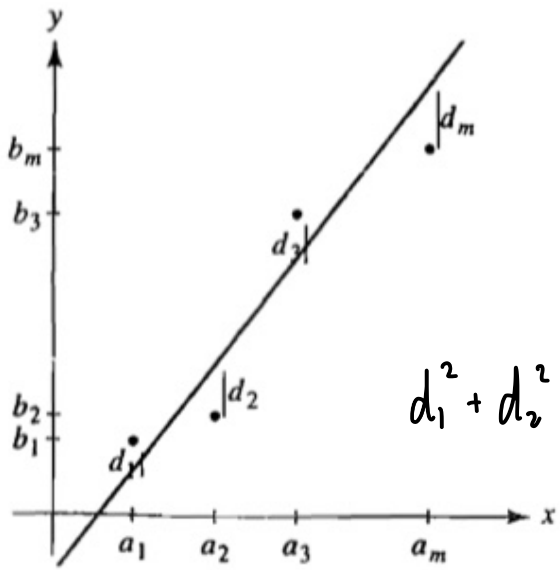
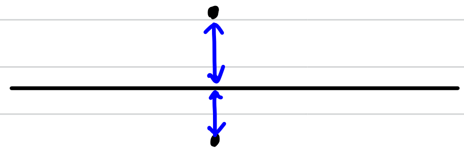


FIGURE 6.13
The distances d_i.



$$d_1^2 + d_2^2 + d_3^2 + \cdots + d_m^2 = \min$$

$L(x) = r_0 + r_1 x$

weight	2	4	5	6
length	6.1	9.1	10.6	12.1

$\Rightarrow L(x) = 1.5x + 3.1$

$L(x) = r_0 + r_1 x$

weight	2	4	5	6	$\leftarrow \vec{a}$
length	6.5	8.5	11.0	12.5	$\leftarrow \vec{b}$

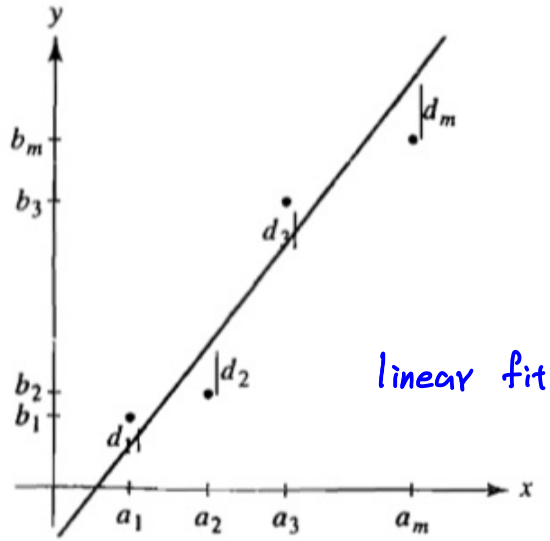
$6.5 \approx r_0 + r_1 \cdot 2$

$8.5 \approx r_0 + r_1 \cdot 4$

$11.0 \approx r_0 + r_1 \cdot 5$

$12.5 \approx r_0 + r_1 \cdot 6$

$$\Rightarrow \underbrace{\begin{bmatrix} 6.5 \\ 8.5 \\ 11.0 \\ 12.5 \end{bmatrix}}_{\vec{b}} \approx \underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} r_0 \\ r_1 \end{bmatrix}}_{\vec{r}}$$



linear fit

FIGURE 6.13
The distances d_i.

$d_1 = 6.5 - (r_0 + r_1 \cdot 2)$

$d_2 = 8.5 - (r_0 + r_1 \cdot 4)$

$d_3 = 11.0 - (r_0 + r_1 \cdot 5)$

$d_4 = 12.5 - (r_0 + r_1 \cdot 6)$

$\Rightarrow \vec{d} = \vec{b} - A\vec{r}$

$d_1^2 + d_2^2 + d_3^2 + d_4^2 = |\vec{d}|^2 = |\vec{b} - A\vec{r}|^2$

$\therefore$ find $\vec{r}$ s.t. $|\vec{b} - A\vec{r}|^2$ is min

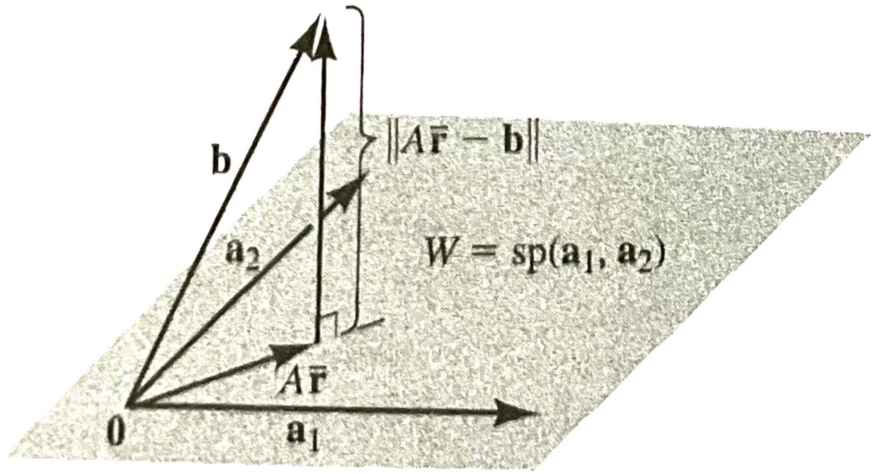


FIGURE 6.14
The length $\|A\bar{r} - b\|$.

$$A\vec{r} = \vec{b}_w = A(A^TA)^{-1}A^T\vec{b}$$

① $\therefore \vec{r} = (A^TA)^{-1}A^T\vec{b}$

② or solve $(A^TA)\underbrace{\vec{x}}_{=\vec{r}} = A^T\vec{b}$

$\boxed{\text{解 } A\vec{x} = \vec{b}}$

weight	2	4	5	6
length	6.5	8.5	11.0	12.5

$$\Rightarrow \underbrace{\begin{bmatrix} 6.5 \\ 8.5 \\ 11.0 \\ 12.5 \end{bmatrix}}_{\vec{b}} \approx \underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} r_0 \\ r_1 \end{bmatrix}}_{\vec{r}}$$

$$A^TA = \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}, \quad (A^TA)^{-1} = \frac{1}{35}\begin{bmatrix} 81 & -17 \\ -17 & 4 \end{bmatrix}$$

$$\therefore \vec{r} = (A^TA)^{-1}A^T\vec{b}$$

$$= \frac{1}{35}\begin{bmatrix} 81 & -17 \\ -17 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 6.5 \\ 8.5 \\ 11.0 \\ 12.5 \end{bmatrix}$$

$$= \frac{1}{35}\begin{bmatrix} 109.5 \\ 53.5 \end{bmatrix} \approx \begin{bmatrix} 3.1 \\ 1.5 \end{bmatrix}$$

Answer: $y = f(x) = 1.5x + 3.1$

ex2, ex3

At a recent boat show, the observations listed in Table 6.2 were made relating the prices b_i of sailboats and their weights a_i. Plotting the data points (a_i, b_i), as shown in Figure 6.12, we might expect a quadratic function of the form

$$y = f(x) = r_0 + r_1x + r_2x^2$$

to fit the data fairly well. ■

$\mathbf{a}_i$ = Weight in tons	2	4	5	8
$\mathbf{b}_i$ = Price in units of \$10,000	1	3	5	12

TABLE 6.6

a_i	b_i	$f(a_i)$
2	1	.959
4	3	3.17
5	5	4.83
8	12	12.0

SOLUTION We write the data in the form $\mathbf{b} \approx A\mathbf{r}$, where A has the form of matrix (7):

$$\begin{bmatrix} 1 \\ 3 \\ 5 \\ 12 \end{bmatrix} \approx \begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}. = \begin{bmatrix} r_0 + r_1 \cdot 2 + r_2 \cdot 2^2 \\ r_0 + r_1 \cdot 4 + r_2 \cdot 4^2 \\ r_0 + r_1 \cdot 5 + r_2 \cdot 5^2 \\ r_0 + r_1 \cdot 8 + r_2 \cdot 8^2 \end{bmatrix}$$

Entering A and $\mathbf{b}$ in either LINTEK or MATLAB, we find that

$$A^TA = \begin{bmatrix} 4 & 19 & 109 \\ 19 & 109 & 709 \\ 109 & 709 & 4993 \end{bmatrix} \quad \text{and} \quad A^T\mathbf{b} = \begin{bmatrix} 21 \\ 135 \\ 945 \end{bmatrix}.$$

Solving the linear system $(A^TA)\mathbf{r} = A^T\mathbf{b}$ using either package then yields

$$\bar{\mathbf{r}} \approx \begin{bmatrix} .207 \\ .010 \\ .183 \end{bmatrix}.$$

Thus, the quadratic function that best approximates the data in the least-squares sense is

$$y \approx f(x) = .207 + .01x + .183x^2.$$

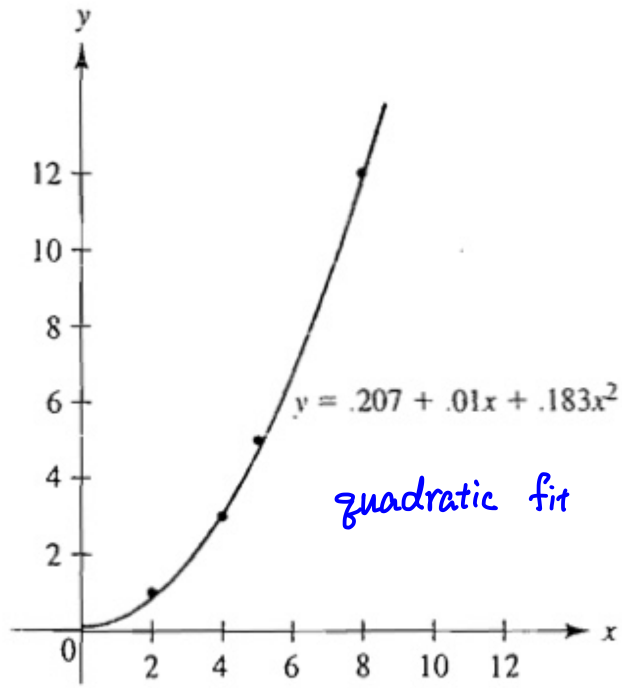


FIGURE 6.17
The graph and data points for Example 3.

$$\begin{array}{c|c|c|c}
a_1 & a_2 & \cdots & a_m \\ \hline
b_1 & b_2 & & b_m
\end{array}$$

$$y = r_0 + r_1 x + r_2 x^2 + r_3 x^3 + \cdots + r_n x^n$$

Data (a_i, b_i) $\Rightarrow$ $b_i \simeq r_0 + r_1 a_i + r_2 a_i^2 + r_3 a_i^3 + \cdots + r_n a_i^n$

$$A = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^n \\ 1 & a_2 & a_2^2 & \cdots & a_2^n \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & a_m & a_m^2 & \cdots & a_m^n \end{bmatrix}.$$

$$\vec{r} = (A^T A)^{-1} A^T \vec{b}$$

or

sol $(A^T A)\vec{r} = A^T \vec{b}$

A population of rabbits on a large island was estimated each year from 1991 to 1994, giving the data in Table 6.3. Knowing that population growth is exponential in the absence of disease, predators, famine, and so on, we expect an exponential function

$$y = f(x) = r e^{sx}$$

TABLE 6.3

$\mathbf{a}_i$ = (Year observed) − 1990	1	2	3	4
$\mathbf{b}_i$ = Number of rabbits in units of 1000	3	4.5	8	17

a_i	1	2	3	4
b_i	3	4.5	8	17

TABLE 6.4

$x = a_i$	1	2	3	4
$y = b_i$	3	4.5	8	17
$z = \ln(b_i)$	1.10	1.50	2.08	2.83

$$y = re^{sx}$$

$$\ln y = \ln(re^{sx}) = \ln r + \ln(e^{sx}) = \ln r + sx$$

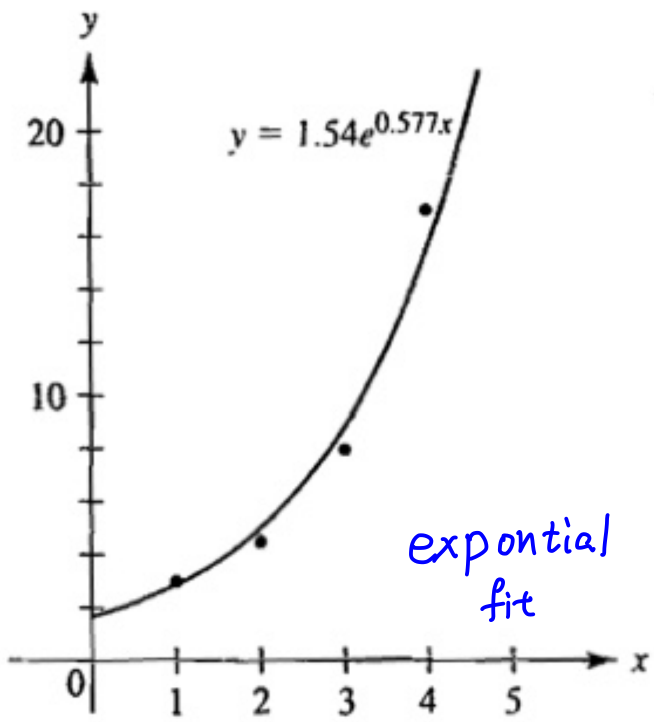
TABLE 6.5

a_i	b_i	$f(a_i)$
1	3	2.7
2	4.5	4.9
3	8	8.7
4	17	15.5

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$$

$$(A^TA)^{-1} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{5} \end{bmatrix}$$

$$(A^TA)^{-1}A^T = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{3}{10} & -\frac{1}{10} & \frac{1}{10} & \frac{3}{10} \end{bmatrix}.$$



$$\vec{r} = (A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} 1 & \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{3}{10} & -\frac{1}{10} & \frac{1}{10} & \frac{3}{10} \end{bmatrix} \begin{bmatrix} 1.10 \\ 1.50 \\ 2.08 \\ 2.83 \end{bmatrix} = \begin{bmatrix} \ln r \\ s \end{bmatrix} \approx \begin{bmatrix} .435 \\ .577 \end{bmatrix}.$$

$e^{0.435} \simeq 1.54 \quad \therefore\ y = 1.54\, e^{0.577x}$

$$\vec{r} = (A^TA)^{-1}A^T\vec{b}$$

or

$$\text{sol } (A^TA)\vec{r} = A^T\vec{b}$$

$$A = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^n \\ 1 & a_2 & a_2^2 & \cdots & a_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_m & a_m^2 & \cdots & a_m^n \end{bmatrix}.$$

weight	2	4	5	6
length	6.5	8.5	11.0	12.5

linear fit

$$A = \begin{array}{c} \begin{array}{cc} \vec{e} & \vec{a} \\ \downarrow & \downarrow \end{array} \\ \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_k \end{bmatrix} \end{array}, \quad A^TA = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_k \end{bmatrix} \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_k \end{bmatrix} = \begin{bmatrix} k & \vec{e}\cdot\vec{a} \\ \vec{e}\cdot\vec{a} & \vec{a}\cdot\vec{a} \end{bmatrix}$$

$$\text{If } \vec{e} \perp \vec{a} \Rightarrow A^TA = \begin{bmatrix} k & 0 \\ 0 & |\vec{a}|^2 \end{bmatrix} \Rightarrow (A^TA)^{-1} = \begin{bmatrix} 1/k & 0 \\ 0 & 1/|\vec{a}|^2 \end{bmatrix}$$

$$\vec{e}\cdot\vec{a} = \sum_i a_i$$

If $A = QR$, with $Q^TQ = I$

$$\vec{r} = (A^TA)^{-1}A^T\vec{b} = \left((QR)^TQR\right)^{-1}(QR)^T\vec{b} = (R^TQ^TQR)^{-1}R^TQ^T\vec{b} = (R^TR)^{-1}R^TQ^T\vec{b}$$

$$= R^{-1}(R^T)^{-1}R^TQ^T\vec{b} = R^{-1}Q^T\vec{b} \Rightarrow R\vec{r} = Q^T\vec{b}$$

EXAMPLE 6 Find the least-squares linear fit of the data points $(-3, 8)$, $(-1, 5)$, $(1, 3)$, and $(3, 0)$.

SOLUTION The matrix A is given by

$$A = \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}.$$

We can see why the symmetry of the x-values about zero causes the column vectors of this matrix to be orthogonal. We find that

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 20 \end{bmatrix}.$$

Then

$$\begin{aligned} \bar{\mathbf{r}} = \begin{bmatrix} r_0 \\ r_1 \end{bmatrix} = (A^T A)^{-1} A^T \mathbf{b} &= \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{20} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 8 \\ 5 \\ 3 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{20} \end{bmatrix} \begin{bmatrix} 16 \\ -26 \end{bmatrix} = \begin{bmatrix} 4 \\ -1.3 \end{bmatrix}. \end{aligned}$$

Thus, the least-squares linear fit is given by $y = 4 - 1.3x$. ∎

△ overdetermined systems of linear equation

$\boxed{\text{solve} \quad A\vec{x}=\vec{b}}$ $\qquad A: m\times n$ matrix

$$\begin{cases} x+y+z=3 \\ 2x-y+3z=4 \end{cases}$$

$$\Rightarrow \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \end{bmatrix}}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$A: 2\times 3$

※eq. ※unknown

※影片錯了,改一下

1-4 29. Mark each of the following True or False.

F a. Every linear system with the same number of equations as unknowns has a unique solution. $m=n$

F b. Every linear system with the same number of equations as unknowns has at least one solution. $m=n$

T c. A linear system with more equations than unknowns may have an infinite number of solutions. $m<n$

T d. A linear system with fewer equations than unknowns may have no solution. $m>n$

overdetermined systems : $m>n$

$$\underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & -1 \end{bmatrix}}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}$$

$A: 5\times 3$

$A\vec{x} \approx \vec{b}$

ex: $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}$

$$A \cdot \vec{x} = \vec{b}$$

ex: $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$

and $\begin{bmatrix} 0 & 2 & 1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

① $\therefore \vec{r} = (A^T A)^{-1} A^T \vec{b} \Rightarrow A\vec{r} \simeq \vec{b}$

② sol: $(A^T A)\vec{r} = A^T \vec{b}$

$$|\vec{b} - A\vec{r}| \text{ min}$$

$$\underbrace{\begin{bmatrix} 1 & 2 & 1 & 0 & -1 \\ 0 & 1 & 3 & 2 & 2 \\ 1 & 1 & 0 & 1 & -1 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & -1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\vec{r}} = \underbrace{\begin{bmatrix} 1 & 2 & 1 & 0 & -1 \\ 0 & 1 & 3 & 2 & 2 \\ 1 & 1 & 0 & 1 & -1 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}}_{\vec{b}}$$

or

$$\begin{bmatrix} 7 & 3 & 4 \\ 3 & 18 & 1 \\ 4 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 3 \end{bmatrix}, \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \approx \begin{bmatrix} -.614 \\ .421 \\ 1.259 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.6447 \\ 0.4518 \\ 0.6497 \\ 2.1015 \\ 0.1980 \end{bmatrix}$$