7.8 An Application to Least Squares

Eg. *Given points* $(1,1)$, $(3,2)$, $(4,3)$ *and functions* 1, x. *Consider* $f(x) = a_0 + a_1 x$.

We want a line $f(x)$ *passing through all the points.*

But the system $\begin{cases} a_0 + a_1 = 1 \\ a_0 + 3a_1 = 2 \\ a_0 + 4a_1 = 3 \end{cases}$ *has no solution.*

Still, there exists a "best-fitting" line for the points.

Eg. $(x_1, y_1),\ (x_2, y_2),\ \cdots, (x_n, y_n) \qquad 1,\ x,\ x^2, \cdots, x^m$
We want to determine a "best-fitting" polynomial of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

for the points.

Eg. *We may consider more general functions than polynomials.*

$$(-1, 0),\ (0, 1),\ (1, 4) \qquad x,\ 2^x$$

We want to determine a "best-fitting" function of the form

$$f(x) = a_0 x + a_1 2^x$$

for the points.

Suppose that

$$(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$$

are given and suppose that $m+1$ functions

$$f_0(x), f_1(x), \cdots, f_m(x)$$

are specified.
Consider $f(x) = a_0 f_0(x) + a_1 f_1(x) + \cdots + a_m f_m(x) \quad (a_i \in \mathbf{R})$.

Problem (in naive form) : Determine a_i's so that $f(x)$ may be the "best- fitting" function for the points $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$.

Let $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ and $M = \begin{bmatrix} f_0(x_1) & f_1(x_1) & \cdots & f_m(x_1) \\ f_0(x_2) & f_1(x_2) & \cdots & f_m(x_2) \\ \vdots & \vdots & & \vdots \\ f_0(x_n) & f_1(x_n) & \cdots & f_m(x_n) \end{bmatrix}$.

$$\begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f_(x_n) \end{bmatrix} = \begin{bmatrix} a_0 f_0(x_1) + a_1 f_1(x_1) + \cdots + a_m f_m(x_1) \\ a_0 f_0(x_2) + a_1 f_1(x_2) + \cdots + a_m f_m(x_2) \\ \vdots \\ a_0 f_0(x_n) + a_1 f_1(x_n) + \cdots + a_m f_m(x_n) \end{bmatrix} = M \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}$$

$$\begin{bmatrix} y_1 - f(x_1) \\ y_2 - f(x_2) \\ \vdots \\ y_n - f(x_n) \end{bmatrix} = Y - M \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}.$$

Problem : Choose $Z = \begin{bmatrix} a_0 & a_1 & \cdots & a_m \end{bmatrix}^T$ such that

$$\|Y - MZ\|^2 = [y_1 - f(x_1)]^2 + [Y_2 - f(x)_2]^2 + \cdots + [y_n - f(x_n)]^2$$

is as small as possible.

A function satisfying the above condition is called a least squares approximating function of the form $f(x) = a_0 f_0(x) + a_1 f_1(x) + \cdots + a_m f_m(x)$.

Write $U = \{MX | X \in \mathbf{R}^m\}$. Then MZ is the closest vector in U to the point Y, so $MZ = \mathrm{proj}_U(Y)$. Thus we have for all $X \in \mathbf{R}^m$

$$\begin{aligned} 0 &= (MX) \cdot (Y - MZ) = (MX)^T (Y - MZ) \\ &= X^T M^T (Y - MZ) = X \cdot [M^T (Y - MZ)]. \end{aligned}$$

Since $M^T(Y - MZ)$ is orthogonal to every vector in $\mathbf{R}^m$, $M^T(Y - MZ) = O$ and so

$$M^T MZ = M^T Y.$$

Thm. *If* $Z = \begin{bmatrix} a_0 & a_1 & \cdots & a_m \end{bmatrix}^T$ *is a solution of*

$$M^T MZ = M^T Y,$$

then

$$f(x) = a_0 f_0(x) + a_1 f_1(x) + \cdots + a_m f_m(x)$$

is a least squares approximating function of the given form.

Eg. *Given* $(1,1),(3,2),(4,3)$*, find the least squares approximating function of the form* $a_0 + a_1 x$*.*

Solution. $f_0(x) = 1$, $f_1(x) = x$.

$$M = \begin{bmatrix} f_0(x_1) & f_1(x_1) \\ f_0(x_2) & f_1(x_2) \\ f_0(x_3) & f_1(x_3) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

$$M^T M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 8 & 26 \end{bmatrix}, \quad M^T Y = \begin{bmatrix} 6 \\ 19 \end{bmatrix}$$

$$M^T M Z = M^T Y, \quad \begin{bmatrix} 3 & 8 \\ 8 & 26 \end{bmatrix} Z = \begin{bmatrix} 6 \\ 19 \end{bmatrix}, \quad Z = \frac{1}{14} \begin{bmatrix} 4 \\ 9 \end{bmatrix}$$

Now we have

$$f(x) = \frac{2}{7} + \frac{9}{14}x.$$

Observe that $f(x)$ *passes through* $\left(1, \frac{13}{14}\right), \left(3, \frac{31}{14}\right), \left(4, \frac{40}{14}\right)$.

Eg. *Given* $(-1,0),(0,1),(1,4)$, *find the least squares approximating function of the form* $a_0x + a_1 2^x$.

Solution. $f_0(x) = x$, $f_1(x) = 2^x$.

$$M = \begin{bmatrix} f_0(x_1) & f_0(x_1) \\ f_0(x_2) & f_1(x_2) \\ f_0(x_3) & f_1(x_3) \end{bmatrix} = \begin{bmatrix} -1 & \frac{1}{2} \\ 0 & 1 \\ 1 & 2 \end{bmatrix}$$

$$M^TM = \begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & \frac{21}{4} \end{bmatrix}, \quad M^TY = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$$

$$M^TMZ = M^TY, \quad Z = \frac{1}{11}\begin{bmatrix} 10 \\ 16 \end{bmatrix}$$

$$f(x) = \frac{10}{11}x + \frac{16}{11}2^x.$$

5.4 Linear Transformation

Def. *A function* $T : \mathbf{F}^n \to \mathbf{F}^m$ *is called a linear transformation if it satisfies, for all* X *and* Y *in* $\mathbf{F}^n$ *and for all* $a \in \mathbf{F}$,

$$T(X+Y) = T(X) + T(Y), \qquad T(aX) = aT(X).$$

Observe that $T(O) = O$, $T(-X) = -T(X)$, and

$$T(a_1X_1 + a_2X_2 + \cdots + a_kX_k) = a_1T(X_1) + a_2T(X_2) + \cdots + a_kT($$

Eg. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ -y \end{bmatrix}$ *and* $S\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y+2z \\ x \end{bmatrix}$ *are linear transformations.*

Def. A : $m \times n$ *matrix. Define* $T_A : \mathbf{F}^n \to \mathbf{F}^m$ *by*

$$T_A(X) = AX$$

for $X \in \mathbf{F}^n$. *Then* T_A *is a linear transformation.*

Eg. $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$.

$$B = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}, T_B\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y + 2z \\ x \end{bmatrix}.$$

Thm. *Let* $T : \mathbf{F}^n \to \mathbf{F}^m$ *be a linear transformation. Then* $T = T_A$ *for the unique matrix* A *given by*

$$A = \begin{bmatrix} T(E_1) & T(E_2) & \cdots & T(E_n) \end{bmatrix},$$

where $\{E_1, E_2, \cdots, E_n\}$ *is the standard basis of* $\mathbf{F}^n$*.*

Proof. $X = x_1 E_1 + x_2 E_2 + \cdots + x_n E_n$

$$\begin{aligned} T(X) &= x_1 T(E_1) + x_2 T(E_2) + \cdots + x_n T(E_n) \\ &= \begin{bmatrix} T(E_1) & T(E_2) & \cdots & T(E_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = AX. \end{aligned}$$

□

Eg. *Let* $R_\theta : \mathbf{R}^2 \to \mathbf{R}^2$ *be the rotation through the angle* θ*. Then* R_θ *is a linear transformation.*
Since $R_\theta(E_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ *and* $R_\theta(E_2) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$*,*

$$\begin{bmatrix} R_\theta(E_1) & R_\theta(E_2) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

induces R_θ*.*

Eg. *Given* $N = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$*, define a linear transformation* $P_N : \mathbf{R}^3 \to \mathbf{R}^3$ *by*

$$P_N(X) = X - \frac{X \cdot N}{\|N\|^2} N.$$

Find the matrix of P_N*.*

Solution.

$$P_N(E_1) = E_1 - \frac{E_1 \cdot N}{\|N\|^2} N = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$$

$$P_N(E_2) = E_2 - \frac{E_2 \cdot N}{\|N\|^2} N = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$P_N(E_3) = E_3 - \frac{E_3 \cdot N}{\|N\|^2} N = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$$

$$\begin{bmatrix} P_N(E_1) & P_N(E_2) & P_N(E_3) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

- Composition

  If we have two linear transformations $T : \mathbf{F}^n \longrightarrow \mathbf{F}^m$ and $S : \mathbf{F}^m \longrightarrow \mathbf{F}^k$, then we can define the composite $S \circ T$ of S and T by

  $$(S \circ T)(X) = S(T(X)) \text{ for all } X \in \mathbf{F}^n.$$

  Thm. *Let $T : \mathbf{F}^n \longrightarrow \mathbf{F}^m$ and $S : \mathbf{F}^m \longrightarrow \mathbf{F}^k$ be linear transformations with matrices B and A, respectively. The composite $S \circ T : \mathbf{F}^n \longrightarrow \mathbf{F}^k$ is the linear transformation with matrix AB.*

  Proof. $(S \circ T)(X) = S(T(X)) = S(BX) = (AB)X$ for $X \in \mathbf{F}^n$. $\square$

Eg. *Find the matrix A of reflection in the X axis of $\mathbf{R}^2$ followed by reflection in the line $y = x$.*

Solution *The matrix of reflection in the X axis is* $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, *and the matrix of reflection in the line $y = x$ is* $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. *Hence*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

- Inverse

Thm. *Let T be a linear transformation with matrix A. Then A is invertible if and only if T has an inverse linear transformation. In this case, T^{-1} is the linear transformation with the matrix A^{-1}.*

Proof. If A is invertible, we define T' by $T'(X) = A^{-1}X$. Then $(T' \circ T)(X) = A^{-1}AX = X$ and $(T \circ T')(X) = AA^{-1}X = X$. Hence, $T' = T^{-1}$.

Conversely, if T^{-1} is the inverse with matrix B, then $T \circ T^{-1} = id$ implies $AB = I$, and $T^{-1} \circ T = id$ implies $BA = I$. Thus $B = A^{-1}$. $\square$

Eg. *Let* R_θ *be the rotation through the angle* θ*. Clearly, the inverse is* $R_\theta^{-1} = R_{-\theta}$*, the rotation through* $-\theta$*. Since the matrix of* R_θ *is* $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$*, the matrix of* R_θ^{-1} *is*

$$A^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}.$$

- One-to-one and Onto

Def. *A linear transformation T is one-to-one (resp. onto) if it is one-to-one (resp. onto) as a function.*

Def. *Let $T : \mathbf{F}^n \to \mathbf{F}^m$ be a linear transformation with matrix A. Then we define*

$$\mathrm{im}T = \{T(X) | X \in \mathbf{F}^n\} = \{AX | X \in \mathbf{F}^n\} = \mathrm{im}A,$$

$$\mathrm{ker}T = \{X \in \mathbf{F}^n | T(X) = O\} = \{X \in \mathbf{F}^n | AX = O\} = \mathrm{null}A.$$

Thm. *Let $T : \mathbf{F}^n \to \mathbf{F}^m$ be a linear transformation. Then T is one-to-one (resp. onto) if and only if $\mathrm{ker} T = O$ (resp. $\mathrm{im} T = \mathbf{F}^m$).*

Rmk. *If follows from*

$$n = \dim(\mathrm{null} A) + \dim(\mathrm{im} A)$$

that

$$n = \dim(\mathrm{ker} T) + \dim(\mathrm{im} T).$$