

Linear Algebra for Statistics

Chapter 8

Instructor: Seokho Lee (이석호)

Hankuk University of Foreign Studies

Chapter 8 직교와 투영

8.1 투영

Theorem (정리 8.1 벡터위의 투영)

$\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ 인 두 벡터 사이의 각이 θ 일 때, 벡터 \mathbf{b} 를 \mathbf{a} 방향으로 투영한 그림자벡터는

$$\mathbf{p} = \frac{\mathbf{b}^T \mathbf{a}}{\mathbf{a}^T \mathbf{a}} \mathbf{a} \quad (8.1)$$

이며, 이 그림자벡터 \mathbf{p} 를 $\text{Span}\{\mathbf{a}\}$ 에 대한 \mathbf{b} 의 투영(*projection*)이라고 한다.

투영된 그림자벡터 \mathbf{p} 는 크기와 방향면에서 다음과 같이 주어진다.

- 크기 :

$$\|\mathbf{b}\| \cdot \cos \theta = \|\mathbf{b}\| \frac{\mathbf{b}^T \mathbf{a}}{\|\mathbf{b}\| \|\mathbf{a}\|} = \frac{\mathbf{b}^T \mathbf{a}}{\|\mathbf{a}\|}$$

- 방향 :

$$\frac{\mathbf{a}}{\|\mathbf{a}\|}$$

따라서 투영 \mathbf{p} 는

$$\mathbf{p} = \frac{\mathbf{b}^T \mathbf{a}}{\|\mathbf{a}\|} \times \frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{\mathbf{b}^T \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} = \frac{\mathbf{b}^T \mathbf{a}}{\mathbf{a}^T \mathbf{a}} \mathbf{a} \quad (8.2)$$

이다.

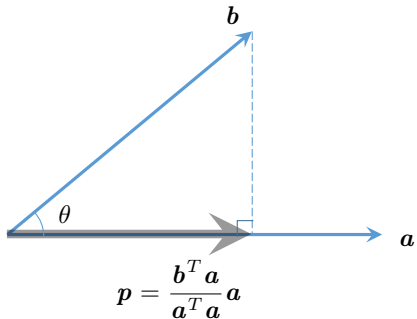


Figure 8.1: 벡터 \mathbf{b} 를 벡터 \mathbf{a} 에 투영시켜 얻은 그림자벡터 \mathbf{p}

Example

아래 벡터 \mathbf{b} 를 $\text{Span}\{\mathbf{a}\}$ 에 투영한 벡터 \mathbf{p} 를 구하고, 좌표평면에 세개의 벡터를 모두 표시하라.

$$\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -5 \\ 0 \end{pmatrix}. \quad (8.3)$$

(풀이) 벡터 \mathbf{p} 는 아래와 같이 얻어진다.

$$\mathbf{p} = \frac{\mathbf{b}^T \mathbf{a}}{\mathbf{a}^T \mathbf{a}} \mathbf{a} = \frac{(2)(-5) + (1)(0)}{2^2 + 1^2} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ -2 \end{pmatrix}. \quad (8.4)$$

세 개의 벡터 그림은 [Figure 8.2](#) 에 도시하였다.

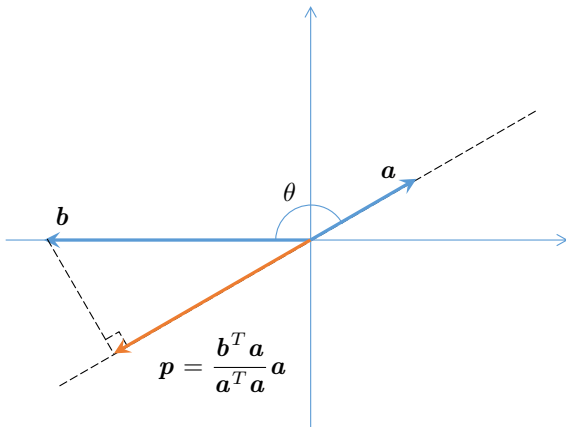


Figure 8.2: 벡터 \mathbf{b} 를 벡터 \mathbf{a} 에 투영시켜 얻은 그림자벡터 \mathbf{p}

Example

이전의 예제에서 벡터 \mathbf{a} 를 $\text{Span}\{\mathbf{b}\}$ 에 투영하여 생긴 벡터 \mathbf{p}' 을 구하고 Figure 8.2 위에 도시하라.

8.2 그램-슈미트 과정(Gram-Schmidt Process)

Definition

벡터공간 V 에 대하여, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ 이 V 의 기저(basis)라고 할 때,

- ① $\mathbf{v}_i^T \mathbf{v}_j = 0$, if $i \neq j$
- ② $\mathbf{v}_i^T \mathbf{v}_j = 1$, if $i = j$

이면 $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ 을 V 의 직교정규기저(orthonormal basis)라 부른다.

Theorem (정리 8.2 그램-슈미트 직교화 과정)

실수공간 \mathbb{R}^n 의 부분공간 V 에 대하여 다음과 같은 과정으로 직교정규기저를 구한다.

- (1) 부분공간 V 를 이루는 기저 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 을 찾는다.
- (2) $\mathbf{w}_1 = \mathbf{a}_1$ 라고 한다.
- (3) $j = 2, 3, \dots, n$ 에 대하여

$$\mathbf{w}_j = \mathbf{a}_j - \left(\frac{\mathbf{a}_j^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{a}_j^T \mathbf{w}_2}{\mathbf{w}_2^T \mathbf{w}_2} \mathbf{w}_2 + \cdots + \frac{\mathbf{a}_j^T \mathbf{w}_{j-1}}{\mathbf{w}_{j-1}^T \mathbf{w}_{j-1}} \mathbf{w}_{j-1} \right) \quad (8.5)$$

을 사용해 직교기저 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ 을 구한다.

- (4) 직교정규기저 $\mathbf{v}_i = \mathbf{w}_i / \|\mathbf{w}_i\|$ ($i = 1, 2, \dots, n$) 를 구한다.

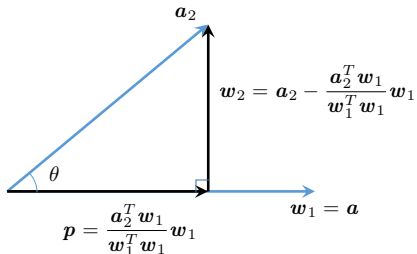


Figure 8.3: 기저 \mathbf{w}_1 과 직교하는 기저 \mathbf{w}_2 를 얻는 그램-슈미트 과정

Figure 8.3 에서 볼 수 있듯이, \mathbf{a}_2 를 \mathbf{w}_1 에 투영한 벡터 \mathbf{p} 와 벡터 \mathbf{a}_2 의 차이 벡터 $\mathbf{a}_2 - \mathbf{p}$ 는 \mathbf{w}_1 과 직교함을 알 수 있다. 이를

$$\mathbf{w}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \mathbf{w}_1 \quad (8.6)$$

라 하면, \mathbf{w}_1 과 \mathbf{w}_2 는 직교한다.

또 다른 벡터 \mathbf{a}_3 가 있다 하자. 이를 이용하여 직교하는 벡터 \mathbf{w}_1 과 \mathbf{w}_2 에 모두 직교하는 \mathbf{w}_3 는 두 벡터 $\mathbf{w}_1, \mathbf{w}_2$ 가 이루는 평면과 직교해야 한다. 이는 \mathbf{a}_3 를 $\mathbf{w}_1, \mathbf{w}_2$ 에 각각 투영시켜 얻을 수 있다. 즉,

$$\mathbf{w}_3 = \mathbf{a}_3 - \left(\frac{\mathbf{a}_3^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{a}_3^T \mathbf{w}_2}{\mathbf{w}_2^T \mathbf{w}_2} \mathbf{w}_2 \right) \quad (8.7)$$

이며, \mathbf{w}_3 는 $\mathbf{w}_1, \mathbf{w}_2$ 와 직교한다.

이를 순차적으로 적용하면

$$\mathbf{w}_j = \mathbf{a}_j - \left(\frac{\mathbf{a}_j^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{a}_j^T \mathbf{w}_2}{\mathbf{w}_2^T \mathbf{w}_2} \mathbf{w}_2 + \cdots + \frac{\mathbf{a}_j^T \mathbf{w}_{j-1}}{\mathbf{w}_{j-1}^T \mathbf{w}_{j-1}} \mathbf{w}_{j-1} \right) \quad (8.8)$$

를 얻게 되며, $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ 은 직교기저이다. 이를 정규화하면 정규직교기저 $\mathbf{v}_i = \mathbf{w}_i / \|\mathbf{w}_i\|$ 를 얻을 수 있다.

Example

세개의 벡터로 생성되는 공간 $S = \text{Span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ 에 대하여 직교정규기저 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 를 구하라.

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}. \quad (8.9)$$

(풀이) 그램-슈미트 과정을 이용하자. 먼저

$$\mathbf{w}_1 = \mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (8.10)$$

이라 두자. 그리고

$$\mathbf{w}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \mathbf{w}_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} - \frac{(1)(1) + (1)(-2) + (1)(1)}{1^2 + 1^2 + 1^2} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \quad (8.11)$$

이며,

Example (continue)

$$\begin{aligned}\mathbf{w}_3 &= \mathbf{a}_3 - \left(\frac{\mathbf{a}_3^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{a}_3^T \mathbf{w}_2}{\mathbf{w}_2^T \mathbf{w}_2} \mathbf{w}_2 \right) \\&= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \frac{(1)(1) + (2)(1) + (3)(1)}{1^2 + 1^2 + 1^2} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{(1)(1) + (2)(-2) + (3)(1)}{1^2 + (-2)^2 + 1^2} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \\&= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\end{aligned}$$

이다. 따라서 $\mathbf{w}_1 = (1, 1, 1)^T$, $\mathbf{w}_2 = (1, -2, 1)^T$, $\mathbf{w}_3 = (-1, 0, 1)^T$ 이며, 이를 정규화하면

$$\mathbf{v}_1 = \frac{1}{\sqrt{3}}(1, 1, 1)^T, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}}(1, -2, 1)^T, \quad \mathbf{v}_3 = \frac{1}{\sqrt{2}}(-1, 0, 1)^T$$

이다.

Example

다음 벡터와 동일한 공간을 생성시키는 직교정규기저를 구하라.

$$\mathbf{a}_1 = (1, 0)^T, \quad \mathbf{a}_2 = (1, 1)^T, \quad \mathbf{a}_3 = (1, 2)^T. \quad (8.12)$$

8.3 투영행렬(Projection Matrix)

Definition

실수공간 \mathbb{R}^n 의 부분공간 V 를 $V = \text{Span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ 라고 정의하고, 행렬 A 가

$$A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k) \quad (8.13)$$

라면 $V = C(A)$ 가 되고

$$P = A(A^T A)^{-1} A^T \quad (8.14)$$

는 $C(A)$ 에 대한 투영행렬(projection matrix)이라 한다.

Remark 행렬 P 가 부분공간 V 의 투영행렬이면 임의의 벡터 \mathbf{x} 에 대해 $P\mathbf{x}$ 는 \mathbf{x} 의 V 로의 투영(projection), 혹은 그림자벡터라고 한다.

Remark 행렬 P 가 부분공간 V 의 투영행렬이면, $\mathbf{x} \in V$ 에 대해 $P\mathbf{x} = \mathbf{x}$ 가 항상 성립한다.

Remark 행렬 A 가 2×2 행렬일 때, A 의 역행렬 A^{-1} 은 다음과 같이 계산된다.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \longrightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad (8.15)$$

Example

두 벡터 \mathbf{a}, \mathbf{b} 로 생성되는 평면에 대한 투영행렬을 구하라.

$$\mathbf{a} = (2, 0)^T, \quad \mathbf{b} = (0, 1)^T. \quad (8.16)$$

Example (continue)

(풀이) \mathbf{a}, \mathbf{b} 를 열로 하는 행렬 A 를

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad (8.17)$$

라고 정의하면, $\text{Span}\{\mathbf{a}, \mathbf{b}\} = C(A)$ 가 된다. 투영행렬은 다음과 같이 얻을 수 있다. 먼저,

$$A^T A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \quad (8.18)$$

이므로 $C(A)$ 에 대한 투영행렬은

$$P = A(A^T A)^{-1} A^T \quad (8.19)$$

$$\begin{aligned} &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \quad (8.20)$$

Example (continue)

이 경우, \mathbb{R}^2 상의 임의의 벡터 \mathbf{x} 는 $C(A) = \text{Span}\{\mathbf{a}, \mathbf{b}\}$ 에 속하기 때문에 $P\mathbf{x} = \mathbf{x}$ 가 성립한다. (실제로 $C(A) = \text{Span}\{\mathbf{a}, \mathbf{b}\} = \mathbb{R}^2$ 이다.)

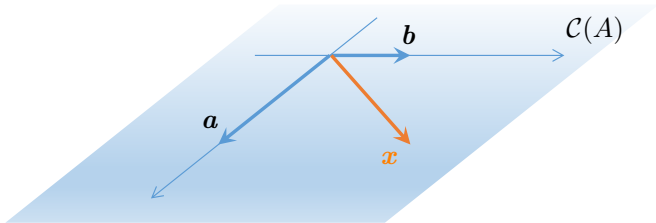


Figure 8.4: 행렬 A 에 대한 $C(A)$ 와 투영행렬 P 에 의한 \mathbf{x} 의 투영 $P\mathbf{x} = \mathbf{x}$

Example

두 벡터 \mathbf{a}, \mathbf{b} 로 생성되는 평면 위에 벡터 \mathbf{c} 의 투영을 구하라.

$$\mathbf{a} = (1, 0, 1)^T, \quad \mathbf{b} = (1, 2, 1)^T, \quad \mathbf{c} = (1, 1, 3)^T. \quad (8.21)$$

(풀이) 두 벡터 \mathbf{a}, \mathbf{b} 를 열벡터로 하는 행렬 A 는 다음과 같다.

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}. \quad (8.22)$$

$A^T A$ 의 역행렬은 다음과 같다.

$$(A^T A)^{-1} = \begin{pmatrix} 2 & 2 \\ 2 & 6 \end{pmatrix}^{-1} = \frac{1}{4} \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}. \quad (8.23)$$

따라서 투영행렬은

Example (continue)

$$P = A(A^T A)^{-1} A^T \quad (8.24)$$

$$\begin{aligned} &= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}. \end{aligned} \quad (8.25)$$

따라서 벡터 \mathbf{c} 의 투영은

$$P\mathbf{c} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}. \quad (8.26)$$

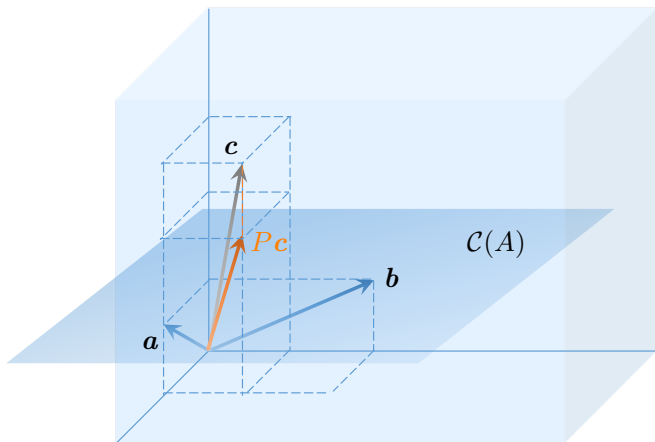


Figure 8.5: 행렬 A 에 대한 $C(A)$ 와 투영행렬 P 에 의한 \mathbf{c} 의 투영 $P\mathbf{c}$

Example

다음 행렬 A 의 열공간 $\mathcal{C}(A)$ 에 대하여, $\mathcal{C}(A)$ 를 생성하는 직교정규기저를 그램-슈미트 과정으로 구하고, 그 기저를 사용하여 투영행렬을 구하여 벡터 \mathbf{c} 의 $\mathcal{C}(A)$ 로의 투영을 구하라.

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}. \quad (8.27)$$

8.4 투영행렬의 성질

Theorem (정리 8.3 대칭-멱등행렬)

투영행렬 $P = A(A^T A)^{-1} A^T$ 는 다음의 성질을 갖는다.

- (1) $P^T = P$. 즉, P 는 대칭(symmetric) 행렬이다.
- (2) $PP = P$. 즉, P 는 멱등(idempotent) 행렬이다.

Proof.

11장에서 다룬 Proposition 11.8에 의해 역행렬이 존재하는 임의의 행렬 M 에 대하여 $(M^T)^{-1} = (M^{-1})^T$ 를 이용하면,

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

이므로 대칭행렬이다. 또한

$$PP = A(A^T A)^{-1} A^T \cdot A(A^T A)^{-1} A^T = A \underbrace{\{(A^T A)^{-1} A^T A\}}_{=I} (A^T A)^{-1} A^T = A(A^T A)^{-1} A^T = P$$

이므로 멱등행렬이다. □

Example

아래는 투영행렬의 예이다. 이들이 대칭-멱등행렬임을 확인하라.

$$P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}. \quad (8.28)$$

8.5 최소제곱법(Least Squares Method)

투영행렬은 회귀분석에서 중요한 역할을 한다.

다음의 입시자료를 고려하자. 6명의 입학성적(x_1 =수능성적, x_2 =학생부 성적)과 이들의 대학재학시 평균학점(y =GPA) 자료를 얻었다고 하자.

x_1 (수능)	x_2 (학생부)	y (GPA)
455	94	3.5
424	87	3.0
502	92	3.7
458	89	2.9
512	90	3.2
412	88	3.4

재학생의 학업성취도가 대학 입학시 성적과 다음의 관계가 있다고 가정하자.

$$y = a_1(\text{수능성적}) + a_2(\text{학생부성적}) = a_1x_1 + a_2x_2 \quad (8.29)$$

따라서, a_1, a_2 값을 정할 수 있으면 수능 및 학생부 성적을 아는 경우 학업성취도를 알 수 있다. 하지만, 모든 학생에 대하여 정확히 위의 식을 만족시키지는 않을 것이므로, 회귀분석에서는 오차(error)항을 고려하여 아래의 모형을 가정한다.

$$y_i = a_1x_{i1} + a_2x_{i2} + \epsilon_i, \quad (i = 1, 2, \dots, 6) \quad (8.30)$$

이를 행렬로 표시하면

$$\underbrace{\begin{pmatrix} 3.5 \\ 3.0 \\ 3.7 \\ 2.9 \\ 3.2 \\ 3.4 \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 455 & 94 \\ 424 & 87 \\ 502 & 92 \\ 458 & 89 \\ 512 & 90 \\ 412 & 88 \end{pmatrix}}_{\mathbf{X}} \underbrace{\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}}_{\mathbf{a}} + \underbrace{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{pmatrix}}_{\boldsymbol{\epsilon}} \quad (8.31)$$

즉, $\mathbf{y} = \mathbf{Xa} + \boldsymbol{\epsilon}$ 라 할 수 있다.

이를 기하학적으로 해석하면 벡터 \mathbf{y} 는 행렬 X 의 열공간 $\mathcal{C}(X)$ 에 속한 벡터인 $X\mathbf{a}$ 와 오차벡터 $\boldsymbol{\epsilon}$ 의 합으로 주어진다.

여기서 \mathbf{a} 값을 결정하는 방법으로 오차항의 길이의 제곱인 $\|\boldsymbol{\epsilon}\|^2 = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon}$ 를 최소로 하는 값으로 결정한다. 이를 최소제곱법(least squares method)이라고 한다.

오차항의 길이를 최소로 하는 것은 결국 벡터 \mathbf{y} 를 $\mathcal{C}(X)$ 에 투영시키는 경우에 얻어진다. 즉, $X\mathbf{a}$ 가 \mathbf{y} 의 $\mathcal{C}(X)$ 로의 투영이 되도록 하는 \mathbf{a} 를 찾음으로 결정한다. \mathbf{y} 의 $\mathcal{C}(X)$ 로의 투영을 $\hat{\mathbf{y}}$ 라고 하고 이때의 \mathbf{a} 값을 $\hat{\mathbf{a}}$ 라고 하면

$$\hat{\mathbf{y}} = P\mathbf{y} = X(X^T X)^{-1} X^T \mathbf{y} = X\hat{\mathbf{a}} \quad (8.32)$$

가 되고, 따라서

$$\hat{\mathbf{a}} = (X^T X)^{-1} X^T \mathbf{y} \quad (8.33)$$

이다. 이를 최소제곱추정량(least squares estimator)이라고 한다.

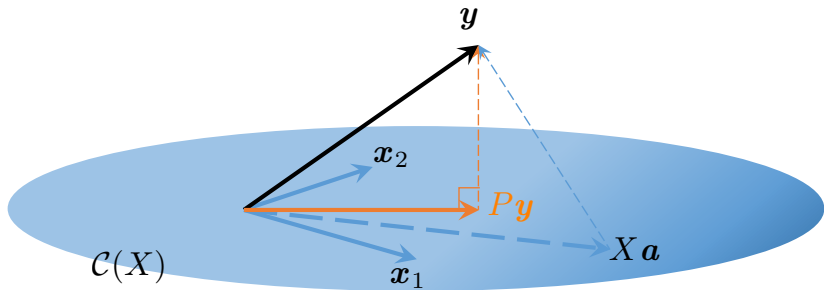


Figure 8.6: $y = Xa + \epsilon$ 에서 최소제곱추정법에 의한 투영 Py

Example

4명의 학생들의 입학성적과 학점평균을 조사하였더니 아래와 같았다. 최소제곱법을 사용하여 학점평균을 수능성적과 학생부성적의 선형결합으로 표현하라.

x_1 (수능)	x_2 (학생부)	y (GPA)
1	1	3.5
1	-1	3.0
2	2	3.0
2	-2	2.5

(풀이) $X^T X$ 의 역행렬은

$$(X^T X)^T = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}^{-1} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \quad (8.34)$$

이므로

Example (continue)

$$\hat{a} = (X^T X)^T X^T \mathbf{y} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 3.5 \\ 3.0 \\ 3.0 \\ 2.5 \end{pmatrix} = \begin{pmatrix} 1.75 \\ 0.15 \end{pmatrix} \quad (8.35)$$

로 최소제곱추정량을 얻는다. 따라서 학점평균과 입학성적과의 관계는

$$\hat{y} = 1.75x_1 + 0.15x_2 \quad (8.36)$$

이다.