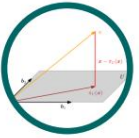


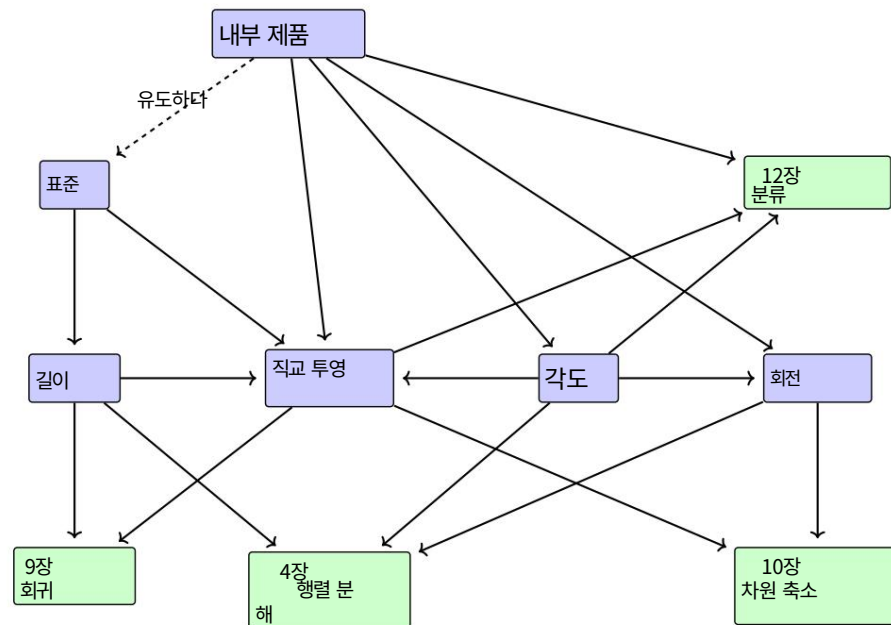
## 해석 기하학



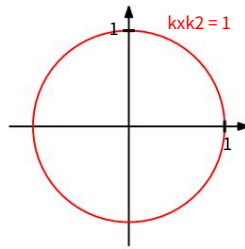
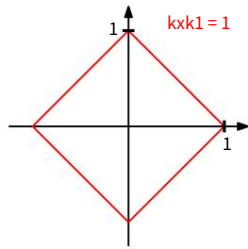
2장에서 우리는 일반적이지만 추상적인 수준에서 벡터, 벡터 공간 및 선형 매핑을 연구했습니다. 이 장에서는 이러한 모든 개념에 기하학적 해석과 직관을 추가할 것입니다. 특히 기하 벡터를 살펴보고 두 벡터 사이의 길이와 거리 또는 각도를 계산합니다. 이를 수행하기 위해 벡터 공간의 기하학을 유도하는 내적을 벡터 공간에 장착합니다. 내적과 그에 상응하는 규범 및 메트릭은 유사성과 거리에 대한 직관적인 개념을 포착하며, 이를 12장에서 지지 벡터 기계를 개발하는 데 사용합니다. 그런 다음 벡터 간의 길이와 각도 개념을 사용하여 직교 투영에 대해 논의합니다. 10장에서 주성분 분석을 논의하고 9장에서 최대우도 추정을 통한 회귀에 대해 논의할 때 중심 역할을 합니다. 그림 3.1은 이 장의 개념이 어떻게 관련되고 책의 다른 장과 어떻게 연결되는지에 대한 개요를 제공합니다.

그림 3.1 이 책에서 소개한 개념의 마인드맵

책의 다른 부분에서 사용되는 경우와 함께 장.



## 3.1 규범



71

그림 3.3  
서로 다른 규범,  
빨간색은  
벡터 세트  
표준 1. 왼쪽:  
맨해튼 규범;  
오른쪽: 유클리드  
거리.

## 3.1 규범

기하학적 벡터를 생각할 때  
원점에서 직관적으로 벡터의 길이는  
원점에서 이 방향 선분의 "끝". 다음에서 우리는  
규범의 개념을 사용하여 벡터의 길이 개념을 논의할 것입니다.

정의 3.1 (규범). 벡터 공간  $V$ 의 노름은 함수입니다.

$$k \cdot k : V \rightarrow \mathbb{R}, \quad (3.1)$$

$$x \mapsto kxk, \quad (3.2)$$

모든  $\lambda \in \mathbb{R}$  길이  $\lambda$ 에 대해 각 벡터  $x$ 의 길이  $kxk \in \mathbb{R}$ 를 할당합니다.  
및  $x, y \in V$  다음 유지:

- 절대적으로 균일:  $k\lambda xk = |\lambda| kxk$
- 삼각형 부등식:  $kx + yk \leq kxk + kyk$
- 양의 정부호:  $kxk > 0$  및  $kxk = 0 \iff x = 0$

기하학적 용어로 삼각형 부등식은 모든 삼각형에 대해,  
두 변의 길이의 합은 반드시 크거나 같아야 합니다.

나머지 변의 길이까지; 그림 3.2를 참조하십시오.

정의 3.1은 일반 벡터 공간  $V$  (섹션 2.4)에 관한 것이지만

이 책에서는 유한 차원 벡터 공간  $\mathbb{R}^n$ 만 고려할 것입니다.

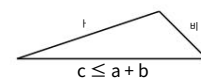
벡터  $x \in \mathbb{R}^n$ 에 대해 다음을 사용하여 벡터의 요소를 나타냅니다.

첨자, 즉  $x_i$ 는  $i$ 입니다.  $x_i$ 는 벡터  $x$ 의 요소입니다.

표준

물론  
동종의  
삼각형 부등식  
양의정의

그림 3.2 삼각형  
불평등.



예 3.1(맨해튼 규범)

$\mathbb{R}^n$ 에 대한 맨해튼 노름은  $x \in \mathbb{R}^n$ 에 대해 다음과 같이 정의됩니다.

$$kxk_1 := \sum_{i=1}^n |x_i|, \quad (3.3)$$

어디  $| \cdot |$  절대값입니다. 그림 3.3의 왼쪽 패널은 모든  
 $kxk_1 = 1$ 인 벡터  $x \in \mathbb{R}^2$ . 맨해튼 노름은  $\ell_1$  노름이라고도 합니다.  
표준.

맨해튼 규범

유클리드 규범

예 3.2(유클리드 규범)  
 $x \in \mathbb{R}^n$ 의 유클리드 노름은 다음과 같이 정의됩니다.

$$\|x\|_2 := \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x^T x} \quad (3.4)$$

유클리드 거리 를 계산하고 원점에서  $x$ 의 유클리드 거리 를 계산합니다. 그림 3.3의 오른쪽 패널은  $\|x\|_2 = 1$ 인 모든 벡터  $x \in \mathbb{R}^2$ 를 보여줍니다. 유클리드 노름은 '2 노름'이라고도 합니다.

`2 규범

주목. 이 책 전체에서 달리 명시되지 않는 한 기본적으로 유클리드 규범(3.4)을 사용합니다. ♦

### 3.2 내부 제곱

내부 제곱은 벡터의 길이와 두 벡터 사이의 각도 또는 거리와 같은 직관적인 기하학적 개념의 도입을 허용합니다. 내적의 주요 목적은 벡터가 서로 직교하는지 여부를 결정하는 것입니다.

#### 3.2.1 내적

우리는 이미 특정 유형의 내적,  $\mathbb{R}^n$ 의 스칼라 곱/내적 에 익숙할 수 있습니다.  
 , 에 의해 주어진

스칼라 곱 내적

$$x \cdot y = \sum_{i=1}^n x_i y_i \quad (3.5)$$

우리는 이 특정 내적을 이 책에서 내적이라고 부를 것입니다. 그러나 내부 제곱은 특정 속성을 가진 보다 일반적인 개념이므로 이제 소개하겠습니다.

#### 3.2.2 일반 내부 제곱

쌍선형 매핑

스칼라를 사용하여 덧셈 및 곱셈과 관련하여 매핑을 재정렬할 수 있는 섹션 2.7의 선형 매핑을 상기하십시오. 양방향 선형 매핑  $\Omega$ 은 두 개의 인수가 있는 매핑이며 각 인수에서 선형입니다. 즉, 벡터 공간  $V$ 를 볼 때 모든  $x, y, z \in V, \lambda, \psi \in \mathbb{R}$ 에 대해 다음을 유지합니다.

$$\Omega(\lambda x + \psi y, z) = \lambda \Omega(x, z) + \psi \Omega(y, z) \quad \Omega(x, \lambda y + \psi z) = \lambda \Omega(x, y) + \psi \Omega(x, z). \quad (3.6)$$

$$(3.7)$$

여기서 (3.6)은 첫 번째 인수에서  $\Omega$ 이 선형임을 주장하고 (3.7)은 두 번째 인수에서  $\Omega$ 이 선형임을 주장합니다 ((2.87) 참조).

정의 3.2.  $V$  를 벡터 공간이라고 하고  $\Omega : V \times V \rightarrow \mathbb{R}$  을 두 벡터를 가져와 실수에 매핑하는 쌍선형 맵이라고 합니다. 그 다음에

- $\Omega$  은 모든  $x, y \in V$  에 대해  $\Omega(x, y) = \Omega(y, x)$  인 경우 대칭이라고 합니다. 즉, 인수의 대칭 순서는 중요하지 않습니다.  $\Omega$  은 다음과 같은 경우 양의 정부호라고 합니다.

$$\forall x \in V \setminus \{0\} : \Omega(x, x) > 0, \Omega(0, 0) = 0. \quad (3.8)$$

양의정의

정의 3.3.  $V$  를 벡터 공간이라고 하고  $\Omega : V \times V \rightarrow \mathbb{R}$  을 두 벡터를 가져와 실수에 매핑하는 쌍선형 맵이라고 합니다. 그 다음에

- 양의 정부호 대칭 쌍선형 맵  $\Omega : V \times V \rightarrow \mathbb{R}$  은  $V$  에 대한 내적이라고 합니다. 일반적으로  $\Omega(x, y)$  대신  $\langle x, y \rangle$  를 씁니다.
- 쌍  $(V, \langle \cdot, \cdot \rangle)$  을 내적 공간 또는 (실제) 벡터 공간 내적 과 내적 공간이라고 합니다. (3.5)에 정의된 내적을 사용하면  $(V, \langle \cdot, \cdot \rangle)$  를 유클리드 벡터 공간이라고 합니다.

내부 제곱

내적이 있는 벡터 공간

이 책에서는 이러한 공간을 내부 제곱 공간이라고 합니다.

유클리드 벡터

우주

예제 3.3 (내적이 아닌 내적)

$V = \mathbb{R}^2$  를 고려하십시오. 우리가 정의한다면

$$\langle x, y \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2x_2 y_2 \quad (3.9)$$

$\langle \cdot, \cdot \rangle$  는 내적이지만 내적과는 다릅니다. 증명은 연습이 될 것입니다.

### 3.2.3 대칭, 양의 유한 행렬

양의 정부호 대칭 행렬은 기계 학습에서 중요한 역할을 하며 내적을 통해 정의됩니다. 4.3절에서 우리는 행렬 분해의 맥락에서 양의 정부호 대칭 행렬로 돌아갈 것입니다. 양의 준정부호 대칭 행렬의 개념은 커널 정의의 핵심입니다(12.4절).

내적  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  (정의 3.3 참조)과  $V$  의 순서화된 기저  $B = (b_1, \dots, b_n)$  을 갖는  $n$ 차원 벡터 공간  $V$  를 고려하십시오. 섹션 2.6.1에서 모든 벡터  $x, y \in V$  는 기본 벡터의 선형 조합으로 작성될 수 있으므로  $x = \sum_{i=1}^n \psi_i b_i \in V$  및  $y = \sum_{j=1}^n \lambda_j b_j \in V$  를 얻습니다.

적절한  $\psi_i, \lambda_j \in \mathbb{R}$  에 대한  $\lambda_j b_j \in V$ . 내적의 쌍 선형성으로 인해 모든  $x, y \in V$  에 대해 다음을 유지합니다.

$$\langle x, y \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_j \psi_i \langle b_j, b_i \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_j \psi_i A_{ji} \quad (3.10)$$

여기서  $A_{ij} := \langle b_i, b_j \rangle$  및  $x, y$  는 기저  $B$  에 대한  $x$  와  $y$  의 좌표입니다. 이는 내적  $\langle \cdot, \cdot \rangle$  가  $A$  를 통해 고유하게 결정됨을 의미합니다. 내적의 대칭 또한 다음을 의미합니다.

대칭입니다. 또한 내적의 양의 정부호는 다음을 의미합니다.

$$\forall x \in V \setminus \{0\} : x^T A x > 0. \quad (3.11)$$

정의 3.4 (대칭, 양의 유한 행렬). (3.11)을 만족 하는 대칭 행렬  $A \in \mathbb{R}^{n \times n}$  을 대칭, 양의 정부호 또는 양의 정부호라고 합니다. >만 (3.11)에서 유지되면 A 는 대칭, 양의 준정부호라고 합니다.

대칭, 양의 정부호

양의 정부호 대칭,  
양의 준정부호

예제 3.4(대칭, 양의 유한 행렬)  
행렬을 고려하십시오

$$A_1 = \begin{pmatrix} 9 & 6 \\ 6 & 5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 9 & 6 \\ 6 & 3 \end{pmatrix}. \quad (3.12)$$

$A_1$  은 대칭이므로 양의 정부호입니다.

$$x^T A_1 x = x_1^2 + 12x_1x_2 + 5x_2^2 \quad (3.13a)$$

$$= 9x_1^2 + 12x_1x_2 + 5x_2^2 = (3x_1 + 2x_2)^2 + x_2^2 > 0 \quad (3.13b)$$

모든  $x \in V \setminus \{0\}$ 에 대해, 대조적으로,  $A_2$  는 대칭이지만  $x^T A_2 x = 9x_1^2 + 12x_1x_2 + 3x_2^2$  가 예를 들어  $x = [2, -3]^T$ 의 경우 0보다 작을 수 있기 때문에 양의 정부호가 아닙니다.

$A \in \mathbb{R}^{n \times n}$  이 대칭이고 양의 정부호 이면

$$(x, y) = x^T A y \quad (3.14)$$

$x$  와  $y$  는 B에 대한  $x, y \in V$ 의 좌표 표현인 순서 기반 B에 대한 내적을 정의합니다.

정리 3.5. 실수 값의 유한 차원 벡터 공간  $V$ 에 대해  $h \cdot, \cdot : V \times V \rightarrow \mathbb{R}$ 은  $V$ 의 차수 기저 B인 경우  $n$ , 와 양의 정부호 대칭 행렬  $A$ 가 존재하는 경우에만 내적입니다.  $A \in \mathbb{R}^{n \times n}$

$$(x, y) = x^T A y. \quad (3.15)$$

$A \in \mathbb{R}^{n \times n}$  이 대칭이고 양의 정부호인 경우 다음 속성이 유지됩니다.

- A의 널스페이스(커널)는 모든  $x \neq 0$ 에 대해  $x^T A x > 0$ 이기 때문에 0으로만 구성됩니다. 이는  $x \neq 0$ 인 경우  $Ax \neq 0$ 임을 의미합니다.
- A의 대각선 요소  $a_{ii}$ 는  $a_{ii} = e_i^T A e_i$ 이므로 양수입니다. 여기서  $e_i$ 는  $\mathbb{R}^n$ 에 있는 표준 기저의  $i$ 번째 벡터입니다.

## 3.3 길이와 거리

75

## 3.3 길이와 거리 섹션 3.1에서 우리

는 이미 벡터의 길이를 계산하는 데 사용할 수 있는 규범에 대해 논의했습니다. 내적과 규범은 어떤 내적도 규범을 유도한다는 점에서 밀접하게 관련되어 있습니다.

내부 제곱은 규범을 유도합니다.

$$\|x\|^2 := \langle x, x \rangle \quad (3.16)$$

내적을 사용하여 벡터의 길이를 계산할 수 있도록 자연스럽게. 그러나 모든 규범이 내적에 의해 유도되는 것은 아닙니다. 맨해튼 규범(3.3)은 해당 내적이 없는 규범의 예입니다. 다음에서는 내적에 의해 유도되는 규범에 초점을 맞추고 길이, 거리 및 각도와 같은 기하학적 개념을 소개합니다.

비고 (코시-슈바르츠 부등식). 내적 벡터 공간  $(V, \langle \cdot, \cdot \rangle)$ 에 대해 유도된 노름  $\| \cdot \|$ 는 코시-슈바르츠 부등식을 충족합니다.

코시-슈바르츠 부등식

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (3.17)$$



## 실시예 3.5 (내적을 이용한 벡터의 길이)

기하학에서 우리는 종종 벡터의 길이에 관심이 있습니다. 이제 내적을 사용하여 (3.16)을 사용하여 계산할 수 있습니다.  $x = [1, 1]^T \in \mathbb{R}^2$ 라고 합시다.

내적을 내적으로 사용하면 (3.16)

$$\|x\|^2 = \langle x, x \rangle = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 + 1 = 2 \quad (3.18)$$

$x$ 의 길이와 같습니다. 이제 다른 내부 제곱을 선택하겠습니다.

$$\langle x, y \rangle := x^T y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + x_2 \quad (3.19)$$

벡터의 노름을 계산하면  $x_1$ 과  $x_2$ 가 동일한 부호(그리고  $x_1 x_2 > 0$ )인 경우 이 내적은 내적보다 작은 값을 반환합니다. 그렇지 않으면 내적보다 큰 값을 반환합니다. 이 내부 제곱으로 우리는

$$\langle x, x \rangle = x_1^2 + x_2^2 = 1 + 1 = 2 \Rightarrow \|x\| = \sqrt{2} \quad (3.20)$$

$x$ 는 내적보다 내적에서 "짧다".

정의 3.6 (거리 및 미터법). 내적 공간  $(V, \langle \cdot, \cdot \rangle)$ 을 고려하십시오. 그 다음에

$$d(x, y) := \|x - y\| = \sqrt{\langle x - y, x - y \rangle} \quad (3.21)$$

$x, y \in V$ 에 대해  $x$ 와  $y$  사이의 거리라고 합니다. dot distance product를 inner product로 사용하면 그 거리를 Euclidean distance라고 합니다. 유클리드 거리

매핑

$$d: V \times V \rightarrow \mathbb{R} \quad (3.22)$$

$$(x, y) \mapsto d(x, y) \quad (3.23)$$

미터법

메트릭이라고 합니다.

주목. 벡터의 길이와 유사하게 벡터 사이의 거리는 내부 제곱이 필요하지 않습니다. 규범이면 충분합니다. 우리에게 규범이 있다면 내부 제곱에 의해 유도되는 거리에 따라 다를 수 있습니다. 내부 제곱 선택. ◆

메트릭  $d$  는 다음을 충족합니다.

양의정의

1.  $d$  는 양의 정부호입니다. 즉, 모든  $x, y \in V$  및  $d(x, y) = 0$  에 대해  $d(x, y) > 0$   $0 \iff x = y$ .

대칭 삼각형

2.  $d$  는 대칭입니다. 즉, 모든  $x, y \in V$  에 대해  $d(x, y) = d(y, x)$  입니다.

부등식

3. 삼각 부등식: 모든  $x, y, z \in V$  에 대해  $d(x, z) \leq d(x, y) + d(y, z)$ .

주목. 언뜻보기에 내부 제곱과 메트릭의 속성 목록은 매우 유사해 보입니다. 그러나 정의 3.3 과 정의 3.6을 비교하여  $\langle x, y \rangle$  및  $d(x, y)$  가 반대 방향으로 동작하는 것을 관찰합니다.

매우 유사한  $x$  및  $y$  는 내부 제곱에 대해 큰 값을 생성하고 메트릭에 대한 작은 값입니다. ◆

### 3.4 각도와 직교성

그림 3.4 언제

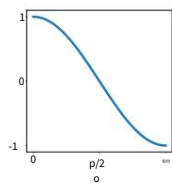
$[0, \pi]$ 로 제한됨

$f(\omega) = \cos(\omega)$

고유한 반환

숫자

간격  $[ -1, 1 ]$ .



벡터의 길이를 정의할 수 있을 뿐만 아니라

두 벡터 사이의 거리, 내적도 기하학을 캡처합니다.

두 벡터 사이의 각도  $\omega$  를 정의하여 벡터 공간의 우리는 사용

두 벡터  $x, y$  사이의 내부 제곱 공간에서 각도  $\omega$  를 정의하기 위한 Cauchy-Schwarz 부등식 (3.17), 이 개념은 우리의

$R^2$  와  $R^3$  의 직관.  $x \neq 0, y \neq 0$  이라고 가정 합니다. 그런 다음

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \|y\|} \leq 1. \quad (3.24)$$

따라서 그림 3.4에서와 같이 고유한  $\omega \in [0, \pi]$  가 존재합니다.

$$\cos \omega = \frac{\langle x, y \rangle}{\|x\| \|y\|}. \quad (3.25)$$

각도

숫자  $\omega$  는 벡터  $x$  와  $y$  사이의 각도입니다. 직관적으로,

두 벡터 사이의 각도는 방향이 얼마나 유사한지를 알려줍니다. 을 위한

예를 들어, 내적을 사용하여  $x$  와  $y$  사이의 각도  $\omega = 4x$ , 즉  $y$

$x$  의 축척된 버전 은 0: 방향이 동일합니다.

## 3.4 각도와 직교성

77

## 예제 3.6(벡터 사이의 각도)

$x = [1, 1]^T \in \mathbb{R}^2$  와  $y = [1, 2]^T \in \mathbb{R}^2$  사이의 각도를 계산해 보겠습니다. 그림 3.5 내적으로 내적을 사용하는 그림 3.5를 참조하십시오. 그런 다음 두 벡터  $x$ ,  $y$  사이의 각도  $\omega$  는 내적을 사용하여 계산됩니다.

$$\cos \omega = \frac{\langle x, y \rangle}{\|x\| \|y\|} = \frac{x^T y}{\sqrt{x^T x} \sqrt{y^T y}} = \frac{2}{3\sqrt{10}}, \quad (3.26)$$

두 벡터 사이의 각도는  $\arccos(\frac{2}{3\sqrt{10}})$  (약  $18^\circ$  )에 해당합니다.  $\approx 0.32 \text{ rad}$ ,

내적의 핵심 기능은 또한 직교인 벡터를 특성화할 수 있다는 것입니다.

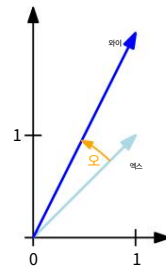
정의 3.7 (직교성). 두 벡터  $x$  와  $y$  는  $\langle x, y \rangle = 0$  인 경우에만 직교하고  $x \perp y$  라고 씁니다. 추가로  $\|x\| = 1$  ,  $\|y\| = 1$  , 즉 벡터가 단위 벡터이면  $x$  와  $y$  는 직교합니다.

직교

이 정의의 의미는 0-벡터가  $y$  에 직교한다는 것입니다.

벡터 공간의 모든 벡터

주목. 직교성은 내적일 필요가 없는 쌍선형 형식에 대한 수직 특이성의 개념을 일반화한 것입니다. 우리의 맥락에서 기하학적으로 직교 벡터는 특정 내적에 대해 직각을 갖는 것으로 생각할 수 있습니다. ♦



## 예제 3.7(직교 벡터)

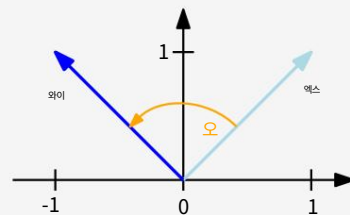


그림 3.6 사이의 각도  $\omega$  두 벡터  $x, y$  는 내부에 따라 변경될 수 있습니다.

제품.

두 벡터  $x = [1, 1]^T, y = [-1, 1]^T \in \mathbb{R}^2$  를 고려하십시오. 그림 3.6을 참조하십시오. 우리는 두 개의 서로 다른 내적을 사용하여 그들 사이의 각도  $\omega$  를 결정하는 데 관심이 있습니다. 내적을 내적으로 사용하면  $x \perp y$  가 되도록  $x$  와  $y$  사이의 각도  $\omega$  가  $90^\circ$  가 됩니다. 그러나 내부 제품을 선택하면

$$\langle x, y \rangle = x^T y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0, \quad (3.27)$$



우리는  $x$  와  $y$  사이의 각도  $\omega$  가 다음과 같이 주어진다 것을 얻습니다.

$$\cos \omega = \frac{\text{점곱}}{\|x\| \|y\|} = \frac{1}{\sqrt{2} \sqrt{2}} \Rightarrow \omega \approx 1.91 \text{ rad} \approx 109.5^\circ, \quad (3.28)$$

$x$  와  $y$  는 직교하지 않습니다. 따라서, 하나의 내적에 대해 직교하는 벡터가 다른 내적에 대해 직교할 필요는 없습니다.

직교 행렬

정의 3.8 (직교 행렬). 정방 행렬  $A \in \mathbb{R}^n \times \mathbb{R}^n$  은 열이 다음과 같이 직교하는 경우에만 직교 행렬입니다.

$$AA^T = I = A^T A, \quad (3.29)$$

그 의미

$$A^{-1} = A^T, \quad (3.30)$$

하는 것이 협약이다  
이러한 행렬을 "직  
교"라고 부르지만  
더 정확한 설명은

즉, 역행렬은 단순히 행렬을 전치하여 얻습니다.

직교 행렬에 의한 변환은 벡터  $x$  의 길이 가 직교 행렬  $A$  를 사용하여 변환할 때 변경되지 않기 때문에 특별합니다. 내적에 대해 다음을 얻습니다.  $(Ax) \cdot (Ax) = x \cdot A^T A x = x \cdot I x = x \cdot x = \|x\|^2$

"정교"가 됩니다.

변환

직교 행렬을 사용  
하면 거리와 각도  
가 유지됩니다.

$$\|Ax\|^2 = x^T A^T A x = x^T I x = x^T x = \|x\|^2. \quad (3.31)$$

더욱이, 두 벡터  $x, y$  사이의 각도는 내적에 의해 측정되며, 직교 행렬  $A$  를 사용하여 두 벡터를 변환할 때도 변경되지 않습니다. 내적을 내적이라고 가정하면 이미지  $Ax$  및  $Ay$  의 각도 다음과 같이 주어진다

$$\cos \omega = \frac{(Ax) \cdot (Ay)}{\|Ax\| \|Ay\|} = \frac{x \cdot A^T A y}{\|x\| \|y\|} = \frac{x \cdot y}{\|x\| \|y\|}, \quad (3.32)$$

정확히  $x$  와  $y$  사이의 각도를 제공합니다. 이것은  $A$  가 있는 직교 행렬  $A$  가 각도 와 거리를 모두 유지한다는 것을 의미합니다. 직교 행렬은 회전(또는 반사)에 대한 변환을 정의합니다. 예제 3.9에서 회

### 3.5 정규직교 기저

2.6.1절에서 우리는 기저 벡터의 속성을 특성화했고  $n$ 차원 벡터 공간에서  $n$ 개의 기저 벡터, 즉 선형 독립인  $n$ 개의 벡터가 필요하다는 것을 발견했습니다. 섹션 3.3과 3.4에서 우리는 벡터의 길이와 벡터 사이의 각도를 계산하기 위해 내적을 사용했습니다. 다음에서 우리는 기저 벡터가 서로 직교하고 각 기저 벡터의 길이가 1인 특별한 경우에 대해 논의할 것입니다. 우리는 이 기저를 직교 기저라고 부를 것입니다.

## 3.6 직교 보완

79

이것을 좀 더 공식적으로 소개해보자.

정의 3.9 (직교정규 기준).  $n$  차원 벡터 공간  $V$  와 기저  $\{b_1, \dots, b_n\}$   $V$ .

$$\langle b_i, b_j \rangle = 0 \text{ for } i \neq j \quad \langle b_i, b_i \rangle = 1 \quad (3.33)$$

$$\langle b_i, b_i \rangle = 1 \quad (3.34)$$

모든  $i$ 에 대해  $j = 1, \dots, n$  이면 기저를 ONB( orthonormal 기저 )라고 합니다. orthonormal 기저 (3.33)만 만족하면 기저를 직교 기저라고 합니다. (3.34)는 모든 기저 벡터가 길이/노름 1을 갖는다는 것을 의미합니다.

온오프

직교 기초

2.6.1절에서 가우스 소거법을 사용하여 벡터 세트에 걸쳐 있는 벡터 공간의 기초를 찾을 수 있음을 상기하십시오. 집합  $\{\tilde{b}_1, \dots, \tilde{b}_n\}$  의 비직교 및 비정규화 기저 벡터. 우리는 그것들을 행렬  $B = [\tilde{b}_1, \dots, \tilde{b}_n]$  및 가우스 제거를 증가된 행렬(섹션 2.3.2)에 적용합니다  $[B \mid B]$  orthonormal basis. orthonormal 기반  $\{b_1, \dots, b_n\}$  은 그람-슈미트 과정 이라고 불린다 (Strang, 2003).

$[B \mid B]$  를 얻으려면

예제 3.8(정규직교 기준)

유클리드 벡터 공간  $\mathbb{R}^n$  에 대한 표준/표준 기저 는 내부 곱이 벡터의 내적인 경우 또는 orthonormal 기저입니다. 벡터

$\mathbb{R}^2$  에서,

$$b_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad b_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (3.35)$$

$b_i$  이후로 직교 기저를 형성

$$\langle b_i, b_j \rangle = 0 \text{ 및 } \|b_1\| = 1 = \|b_2\|.$$

12장과 10장에서 지지 벡터 기계와 주성분 분석에 대해 논의할 때 직교 기저의 개념을 활용할 것입니다.

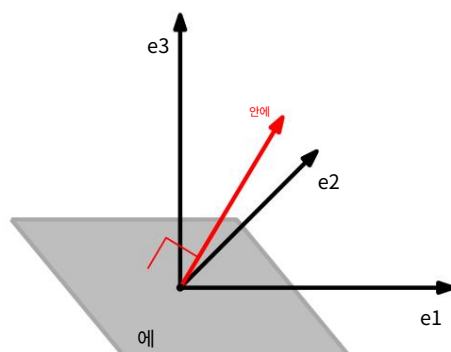
## 3.6 Orthogonal Complement 직교성을

정의했으므로 이제 서로 직교하는 벡터 공간을 살펴보겠습니다. 이것은 기하학적 관점에서 선형 차원 축소를 논의할 10장에서 중요한 역할을 합니다.

$D$  차원 벡터 공간  $V$  와  $M$  차원 하위 를 고려하십시오.

공간  $U \subseteq V$ . 그런 다음  $V$ 의 직교 보수  $U^\perp$  부분 공간은 모든 벡터  $u \in U$  와  $v \in U^\perp$  에 대해  $\langle u, v \rangle = 0$  인 벡터  $v$ 의 집합입니다. 또한  $U \cap U^\perp = \{0\}$  이고  $U \cup U^\perp = V$  이므로  $V = U \oplus U^\perp$  이 됩니다.

그림 3.7 A 평면  
유  
3차원 벡터 공간  
은 직교 보수  $U^\perp$   
에 걸쳐 있는 법선  
벡터로 설명할 수  
있습니다.



로 독특하게 분해

$$\text{중} \quad \text{엑스} = \text{엑스}_{m=1} \lambda_m b_m + \sum_{j=1}^M \psi_j b_j^\perp, \lambda_m, \psi_j \in \mathbb{R}, \quad (3.36)$$

여기서  $(b_1, \dots, b_M)$ 은  $U$ 의 기저이고  $(b_j^\perp)$  따라  $b_1^\perp, \dots, b_M^\perp$ 은  $U^\perp$ 의 기저입니다.

서 직교 보수는 3차원 벡터 공간에서 평면  $U$  (2차원 부분 공간)를 설명하는 데에도 사용  
할 수 있습니다.

보다 구체적으로, 평면  $U$ 에 직교하는  $\|w\| = 1$ 인 벡터  $w$ 는  $U^\perp$ 의 기저 벡터이다. 그림  
3.7은 이 설정을 보여줍니다.  $w$ 에 직교하는 모든 벡터는 (구성상) 평면  $U$ 에 있어야 합니  
다. 벡터  $w$ 는  $U$ 의 법선 벡터라고 합니다.

법선 벡터

일반적으로 직교 보수는  $n$ 차원 벡터 및 아핀 공간에서 초평면을 설명하는 데 사용할 수  
있습니다.

### 3.7 함수의 내적

지금까지 길이, 각도 및 거리를 계산하기 위해 내적의 속성을 살펴보았습니다. 우리는 유한  
차원 벡터의 내적에 초점을 맞췄습니다. 다음에서 우리는 다른 유형의 벡터의 내적의 예를  
살펴볼 것입니다: 함수의 내적.

지금까지 논의한 내적은 유한한 수의 항목이 있는 벡터에 대해 정의되었습니다. 벡터  $x \in \mathbb{R}^n$ 을  $n$ 개의 함수 값을 갖는 함수로 생각할 수 있습니다. 내적의 개념은 무한한 수의 항  
목(가산 무한)과 연속값 함수(무한 무한)를 가진 벡터로 일반화할 수 있습니다. 그런 다음  
벡터의 개별 구성 요소에 대한 합(예: 식 (3.5) 참조)이 적분으로 바뀝니다.

두 함수  $u: \mathbb{R} \rightarrow \mathbb{R}$  및  $v: \mathbb{R} \rightarrow \mathbb{R}$ 의 내적은 다음과 같습니다.  
한정적분으로 정의

$$\langle u, v \rangle := \int_a^b u(x)v(x)dx \quad (3.37)$$

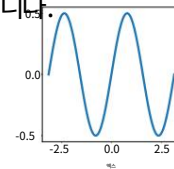
## 3.8 직교 투영

81

하한 및 상한에 대해 각각  $a, b < \infty$ 입니다. 일반적인 내적과 마찬가지로 내적을 보고 규범과 직교성을 정의할 수 있습니다. (3.37)이 0으로 평가되면 함수  $u$  와  $v$  는 직교합니다. 앞의 내적을 수학적으로 정확하게 만들기 위해서는 측정값과 적분의 정의를 처리해야 하며, 이는 힐베르트 공간의 정의로 이어집니다. 또한 유한 차원 벡터의 내적과 달리 함수의 내적은 발산할 수 있습니다(무한 값을 가짐). 이 모든 것은 우리가 이 책에서 다루지 않는 실제 및 기능 분석에 대한 좀 더 복잡한 세부 사항으로 뛰어드는 것을 필요로 합니다.

예제 3.9(함수의 내적)

$u = \sin(x)$  및  $v = \cos(x)$ 를 선택 하면 피적분  $f(x) = u(x)v(x)$  그림 3.8  $f(x) = \sin(x)\cos(x)$  (3.37)은 그림 3.8에 나와 있습니다. 우리는 이 함수가 홀수임을 알 수 있습니다. 즉,  $f(-x) = -f(x)$ 입니다. 따라서  $f(x)$ 의 극한 이  $a = -\pi, b = \pi$  인 적분은 0으로 평가됩니다. 따라서  $\sin$  과  $\cos$  는 직교 함수입니다.



주목. 또한 기능 모음이

$$\{1, \cos(x), \cos(2x), \cos(3x), \dots\} \quad (3.38)$$

$\pi$  에서  $\pi$ 로 적분하면 는 직교합니다. 즉, 함수 쌍이 서로 직교합니다. (3.38)의 함수 모음은  $[-\pi, \pi]$ 에서 짝수이고 주기적인 함수의 큰 부분 공간에 걸쳐 있으며, 이 부분 공간에 함수를 투영하는 것은 푸리에 급수의 기본 아이디어입니다. ♦

섹션 6.4.6에서 우리는 두 번째 유형의 비전통적 내적, 즉 확률 변수의 내적을 살펴볼 것입니다.

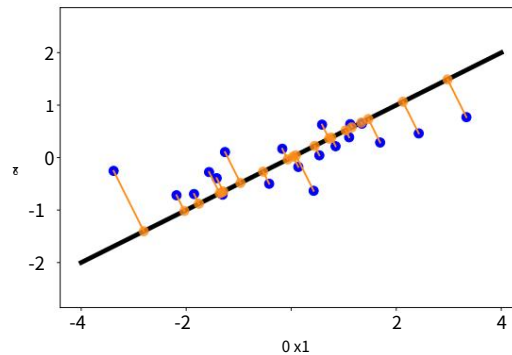
## 3.8 직교 투영

투영은 회전 및 반사 외에 선형 변환의 중요한 클래스이며 그래픽, 코딩 이론, 통계 및 기계 학습에서 중요한 역할을 합니다. 기계 학습에서 우리는 종종 고차원 데이터를 다룹니다. 고차원 데이터는 종종 분석하거나 시각화하기 어렵습니다. 그러나 고차원 데이터는 소수의 차원에만 대부분의 정보가 포함되어 있고 대부분의 다른 차원은 데이터의 주요 속성을 설명하는 데 필수적이지 않다는 속성을 갖고 있는 경우가 많습니다. 고차원 데이터를 압축하거나 시각화하면 정보가 손실됩니다. 이 압축 손실을 최소화하기 위해 데이터에서 가장 유익한 차원을 찾는 것이 이상적입니다. 1장에서 논의한 바와 같이, "특징"은 데이터가 벡터로 표현될 수 있는 일반적인 표현이며, 이 장에서는 데이터 압축을 위한 몇 가지 기본 도구에 대해 논의할 것입니다. 더 구체적으로 말하면, 원래의 고차원 데이터를 저차원 기능 공간에 투영하고 이 저차원 공간에서 작업하여 데이터 세트에 대해 더 배우고 관련 패턴을 추출할 수 있습니다. 예를 들어, 기계

대표.

그림 3.9

직교  
투영(주황색  
점)의  
2차원  
데이터세트(파란색 점)  
에  
일차원적인  
부분공간(직선  
선).



Pearson(1901) 및 Hotelling(1933)의 주성분 분석(PCA) 및 심층 신경망(예: 심층 신경망)과 같은 학습 알고리즘

자동 인코더(Deng et al., 2010))는 차원 축소라는 아이디어를 많이 이용합니다. 다음에서 우리는 직교 투영에 초점을 맞출 것입니다.

선형 차원 축소를 위해 10장에서 사용할 것입니다.

분류에 대해서는 12장에서 참조하십시오. 우리가 논의하는 선형 회귀조차도

9장에서 직교 투영을 사용하여 해석할 수 있습니다. 주어진

저차원 부분 공간, 고차원의 직교 투영

데이터는 가능한 한 많은 정보를 유지하고 차이를 최소화합니다.

원본 데이터와 해당 투영 사이의 오류. 이러한 직교 투영에 대한 설명은 그림 3.9에 나와 있습니다. 전에

우리는 이러한 투영을 얻는 방법을 자세히 설명하고 투영이 무엇인지 정의합니다.

실제로 있습니다.

정의 3.10 (투영).  $V$ 를 벡터 공간이라고 하고  $U \subseteq V$  a

투사

$U$ 의 부분공간 선형 매핑  $\pi: V \rightarrow U$ 는 다음과 같은 경우 투영이라고 합니다.

$$\pi \circ \pi = \pi$$

선형 매핑은 변환 행렬로 표현될 수 있기 때문에(참조

섹션 2.7), 앞의 정의는 특수 종류에도 동일하게 적용됩니다.

투영 행렬

변환 행렬, 다음을 나타내는 투영 행렬  $P$

$$P^2 = P$$

속성  $P$  다음에서 벡터의 직교 투영을 유도합니다.

내적 공간  $(R^n, \langle \cdot, \cdot \rangle)$ 을 부분 공간으로. 선이라고도 하는 1차원 부분공간부터 시작하겠습니다. 달리 언급하지 않으면 내적  $\langle x, y \rangle = x \cdot y$ 를 내적이라고 가정합니다.

선

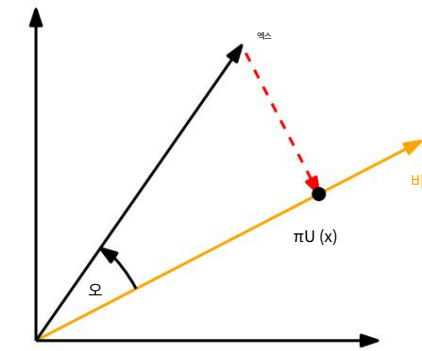
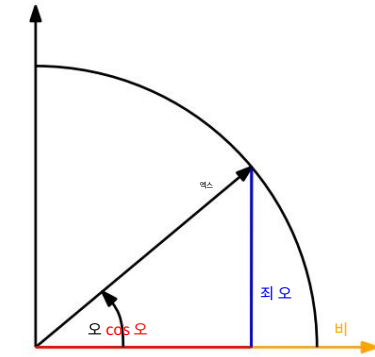
### 3.8.1 1차원 부분공간(선)에 대한 투영

기저 벡터  $b \in R^n$ 을 갖는 원점을 통과하는 선(1차원 부분공간)이 주어졌다고 가정합니다. 선은 1차원 부분공간입니다.

$U \subseteq R^n$ 은  $b$ 로 확장됩니다.  $x \in R^n$ 을  $U$ 에 투영할 때

$x$ 에 가장 가까운 벡터  $\pi_U(x) \in U$ . 기하학적 인수를 사용하여

## 3.8 직교 투영

(a) 기저 벡터  $b$  를 갖는 부분공간  $U$  에  $x \in \mathbb{R}^2$  의 투영.(b)  $\|x\| = 1$  인 2차원 벡터  $x$  를  $b$  에 걸쳐 있는 1차원 부분공간에 투영합니다.

83

그림 3.10 1차원 부분공간에 대한 투영의 예.

우리는 투영  $\pi_U(x)$  의 일부 속성을 특성화합니다 (그림 3.10(a)는 예시로 사용됨).

- 투영  $\pi_U(x)$  는  $x$  에 가장 가깝습니다. 여기서 "가장 가까운"은 거리  $\|x - \pi_U(x)\|$  가 최소임을 의미합니다.  $\pi_U(x)$  에서  $x$  까지 의 세그먼트  $\pi_U(x) - x$  는  $U$  에 직교하므로  $U$  의 기저 벡터  $b$  가 됩니다. 직교성 조건은 벡터 사이의 각도 때문에  $\langle \pi_U(x) - x, b \rangle = 0$  을 산출합니다. 내적을 통해 정의됩니다.
- $U$  에 대한  $x$  의 투영  $\pi_U(x)$  는  $U$  의 요소여야 하며 따라서  $U$  에 걸쳐 있는 기저 벡터  $b$  의 배수여야 합니다. 따라서 일부  $\lambda \in \mathbb{R}$  에 대해  $\pi_U(x) = \lambda b$  입니다.

 $\lambda$  는 다음과 같습니다. $b$  에 대한  $\pi_U(x)$  의 좌표.

다음 세 단계에서 좌표  $\lambda$ , 투영  $\pi_U(x) \in U$  및 투영 행렬  $P$  를 결정합니다.

모든  $x \in \mathbb{R}^n$  을  $U$  에 매핑합니다.

1. 좌표  $\lambda$  찾기. 직교성 조건 산출

$$\langle x - \pi_U(x), b \rangle = 0 \iff \langle x - \lambda b, b \rangle = 0 \quad (3.39)$$

이제 내적의 쌍선형성을 이용하여 다음과 같은 결과에 도달할 수 있습니다.

$$\langle x, b \rangle - \lambda \langle b, b \rangle = 0 \iff \lambda = \frac{\langle x, b \rangle}{\langle b, b \rangle} \quad (3.40)$$

일반적인 내적을 사용하면  $\lambda = \langle x, b \rangle / \|b\|^2$  이면  $b$  를 얻습니다.

마지막 단계에서 우리는 내부 제품이 대칭이라는 사실을 이용했습니다.  $\langle \cdot, \cdot \rangle$  를 내적으로 선택하면

$$\lambda = \frac{\langle x, b \rangle}{\langle b, b \rangle} = \frac{\langle b, x \rangle}{\|b\|^2} \quad (3.41)$$

$\|b\| = 1$  이면 투영의 좌표  $\lambda$  는  $\langle b, x \rangle$  이다.

2. 투영점  $\pi_U(x) \in U$  찾기  $\pi_U(x) = \lambda b$ 이므로 (3.40)

$$\pi_U(x) = \lambda b = b = \frac{\langle x, b \rangle}{\|b\|^2} b, \quad (3.42)$$

여기서 마지막 평등은 내적에 대해서만 유지됩니다. 정의 3.1을 통해 다음과 같이  $\pi_U(x)$ 의 길이를 계산할 수도 있습니다.

$$\|\pi_U(x)\| = \|\lambda b\| = |\lambda| \|b\|. \quad (3.43)$$

따라서 우리의 투영은 길이  $\|b\|$ 의 길이를 곱합니다. 이것은 또한  $\lambda$ 가 1차원 부분공간  $U$ 에 걸쳐 있는 기저 벡터  $b$ 에 대한  $\pi_U(x)$ 의 좌표라는 직관을 추가합니다.

내적을 내적으로 사용하면

$$\|\pi_U(x)\|^2 \stackrel{(3.42)}{=} \left\| \frac{\langle x, b \rangle}{\|b\|^2} b \right\|^2 \stackrel{(3.25)}{=} \frac{\langle x, b \rangle^2}{\|b\|^4} \|b\|^2 = \frac{\langle x, b \rangle^2}{\|b\|^2} \quad (3.44)$$

여기서  $\omega$ 는  $x$ 와  $b$  사이의 각도입니다. 이 방정식은 삼각법에서 익숙해야 합니다.  $\|x\| = 1$ 이면  $x$ 는 단위 원에 있습니다.  $b$ 에 걸쳐 있는 수평 축에 대한 투영은 정확히  $\cos \omega$ 이고 해당 벡터의 길이  $\pi_U(x) = \|\cos \omega\|$ 입니다. 그림 3.10(b)에 그림이 나와 있습니다.

가로축은 1차원

부분 공간.

3. 투영 행렬  $P$  찾기  $\pi_U$ . 우리는 투영이 선형 매핑이라는 것을 알고 있습니다(정의 3.10 참조). 따라서  $\pi_U(x) = P \pi x$ 와 같은 투영이 존재합니다. 내적을 내부 행렬  $P$   $\pi_U$  곱으로 하고

$$\pi_U(x) = \lambda b = b \lambda = b \frac{\langle x, b \rangle}{\|b\|^2} = \frac{\langle x, b \rangle}{\|b\|^2} b, \quad (3.45)$$

우리는 즉시 그것을 본다

$$P = \frac{bb^T}{\|b\|^2}. \quad (3.46)$$

투영 행렬은 항상 대칭입니다.

$bb^T$  (결과적으로  $P$ )는 대칭 행렬(순위 1)이고  $\|b\|^2 = b^T b$ 이고  $b$ 는 스칼라입니다.

투영 행렬  $P$ 는 벡터  $x \in \mathbb{R}^n$ 을  $b$ 방향( $b$ 에 의해 확장된 부분공간  $U$ )의 원점을 통과하는 선에 투영합니다.

주목. 투영  $\pi_U(x) \in \mathbb{R}^n$ 은 여전히  $n$ 차원 벡터이며 스칼라가 아닙니다. 그러나 투영을 나타내기 위해 더 이상  $n$ 좌표가 필요하지 않지만 부분 공간  $U: \lambda$ 에 걸쳐 있는 기저 벡터  $b$ 에 대해 표현하려는 경우 하나의 좌표만 필요합니다.

## 3.8 직교 투영

85

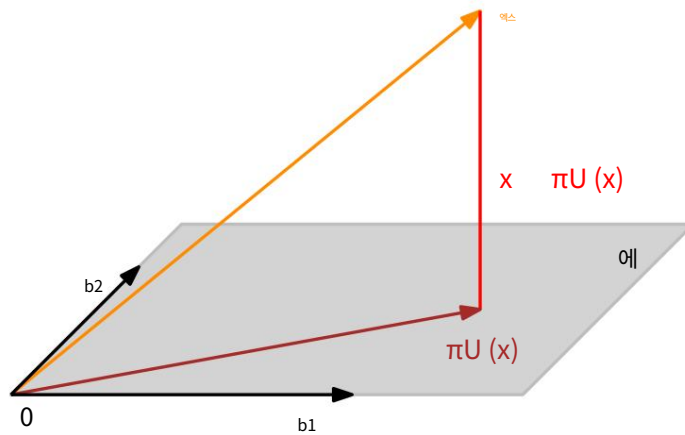


그림 3.11  $b_1, b_2$  를 기저로 하는 2차원 부분공간  $U$ 에 대한 투영.  $U$ 에 대한  $x \in \mathbb{R}^3$ 의 투영  $\pi_U(x)$ 는 선형 조합으로 표현될 수 있습니다.

$b_1, b_2$  및 변위 벡터  $x - \pi_U(x)$ 는  $b_1$ 과  $b_2$  모두에 직교합니다.

$$P = \frac{bb^T}{b^T b} = \frac{1}{9} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}$$

$$\pi_U(x) = P \pi x = \frac{1}{9} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix} \in \text{스팬} \left[ \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right]. \quad (3.48)$$

다. 즉,  $P \pi \pi_U(x) = \pi_U(x)$ 의 적용에  $\pi_U(x)$ 를 사용하면  $\pi_U(x)$ 는  $P$ 에 의해 고정된다.

$P$ 를 만족  $P^2 = P$  모든  $x$ 에 대해  $\pi x = P \pi x$ .

$U$ 가 기저가 아닌 일련의 스페닝 벡터로 주어지면 기저  $b_1, \dots, b_m$  진행하기 전에.



기저 벡터  $b_1, \dots, \pi_U(x) = P_m$  인  $U$  의  $b_m$

가져 벡터  
 $B \in \mathbb{R}^{n \times m}$  의 열을  
 형성합니다. 여기서  $B =$   
 $[b_1, \dots, b_m]$ .

$$\pi_U(x) = \sum_{i=1}^m \lambda_i b_i = B\lambda, \quad$$

$$B = [b_1, \dots, b_m] \in \mathbb{R}^{n \times m}, \lambda = [\lambda_1, \dots, \lambda_m] \in \mathbb{R}^m,$$

$$b_1^T(x - \pi_U(x)) = 0$$

$$b_m^T(x - \pi_U(x)) = 0$$

$$b_1^T(x - B\lambda) = 0$$

$$b_m^T(x - B\lambda) = 0$$

$$b_1^T(x - B\lambda) = 0 \iff B^T(x - B\lambda) = 0$$

$$\iff B^T B \lambda = B^T x.$$

정규 방정식

$$\lambda = (B^T B)^{-1} B^T x.$$

의사 역

## 3.8 직교 투영

87

$B > B$  증가된 수치 안정성 및 양의 정부호도를 보장합니다. 이 "릿지"는 베이저안 추론을 사용하여 엄격하게 파생될 수 있습니다.

2. 투영  $\pi U(x) \in U$ 를 구합니다. 우리는 이미  $\pi U(x) = B\lambda$ . 따라서 (3.57)

$$\pi U(x) = B(B > B)^{-1} B^T x.$$

3. 투영 행렬  $P$  찾기 (3.58)에서  $P \pi x = \pi U(x)$ 를 푸는 투영 행렬은 다음과 같아야 함을 즉시 알 수 있습니다.

$$P = B(B > B)^{-1} B^T.$$

주목. 일반 부분 공간에 투영하기 위한 솔루션은 1D 경우를 특별한 경우로 포함합니다.  $\dim(U) = 1$ 이면  $B > B \in \mathbb{R}$ 은 스칼라이고 투영 행렬을 (3.59) PP로 다시 작성할 수 있습니다.

$$P = \frac{B B^T}{B^T B} \text{ 정확히 (3.46)의 투영 행렬입니다.}$$



부분공간  $U = \text{span} \left[ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right] \subseteq \mathbb{R}^3$  및  $x = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^3$  찾기

dence)  $U$ 의 기저 벡터를 행렬  $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$ 에 씁니다.

둘째, 행렬  $B > B$ 와 벡터  $B^T x$ 를 계산합니다.

$$B > B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}, \quad B^T x = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

셋째, 정규 방정식  $B > B \lambda = B^T x$ 를 풉니다.

$$\begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix} \iff \lambda = \begin{pmatrix} 5 \\ -3 \end{pmatrix}.$$

$$\pi U(x) = B \lambda = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}.$$

투영 오류

투영 오차는 재구성 오차라고도 합니다.

대응하는 투영 오차는 원래 벡터와  $U$ 에 대한 투영 사이의 차이 벡터의 노름입니다. 즉,

$$\|x - \pi_U(x)\|_2 = \sqrt{6}. \quad (3.63)$$

다섯째, 투영 행렬(모든  $x \in \mathbb{R}^3$ 에 대해)은 다음과 같이 주어집니다.

$$\pi_U = B(B^T B)^{-1} B^T = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}. \quad (3.64)$$

결과를 확인하기 위해 (a) 변위 벡터  $\pi_U(x) - x$ 가  $U$ 의 모든 기저 벡터에 직교하는지 확인하고 (b) 다음을 확인할 수 있습니다.

$\pi_U^2 = \pi_U$  (정의 3.10 참조).

주목. 투영  $\pi_U(x)$ 는  $m$ 차원 부분공간  $U \subseteq \mathbb{R}^n$ 에 있지만 여전히  $\mathbb{R}^n$ 의 벡터입니다. 그러나 투영된 벡터를 나타내려면  $m$  좌표  $\lambda_1, \dots, \lambda_m$ 은 기저 벡터  $b_1, \dots, b_m$ 의 선형 조합입니다. ♦

주목. 일반적인 내적이 있는 벡터 공간에서는 ♦ 내적의 수단으로 정의되는 각도와 거리를 계산할 때 주의해야 합니다.

우리는 찾을 수 있어

대략적인 솔루션

투영을 사용하여

풀 수 없는 선형 방정식 시스템.

투영을 통해 해가 없는 선형 시스템  $Ax = b$ 가 있는 상황을 볼 수 있습니다. 이것은  $b$ 가  $A$ 의 범위에 있지 않다는 것을 의미합니다. 즉, 벡터  $b$ 는  $A$ 의 열에 의해 확장되는 부분공간에 있지 않습니다. 선형 방정식을 정확하게 풀 수 없다는 점을 감안할 때 근사 솔루션을 찾을 수 있습니다. 아이디어는  $b$ 에 가장 가까운  $A$ 의 열에 의해 확장되는 부분 공간에서 벡터를 찾는 것입니다. 즉,  $A$ 의 열에 의해 확장되는 부분 공간에 대한  $b$ 의 직교 투영을 계산합니다. 이 문제는 실제로 자주 발생하며, 해를 과결정 시스템의 최소제곱해(내적을 내적으로 가정)라고 합니다. 이것은 섹션 9.4에서 더 논의됩니다. 재구성 오류(3.63)를 사용하는 것은 주성분 분석(10.3절)을 도출하는 한 가지 가능한 접근 방식입니다.

최소제곱해

주목. 우리는 기저 벡터  $\{b_1, \dots, b_k\}$ . 이 기저가 ONB인 경우, 즉 (3.33) 및 (3.34)가 충족되면 투영 방정식 (3.58)은  $\pi_U(x) = BB^T x$ 로 크게 단순화됩니다.

$$B^T B = I \quad (3.65)$$

$B^T B = I$ 가 좌표가 있는  $I$ 이기 때문에

$$\lambda = B^T x. \quad (3.66)$$

이것은 더 이상 (3.58)에서 역을 계산할 필요가 없다는 것을 의미합니다. ♦ 계산 시간을 절약합니다.

## 3.8 직교 투영

89

## 3.8.3 Gram-Schmidt Orthogonalization Projections

는  $n$ 차원 벡터 공간  $V$ 의 모든 기저  $(b_1, \dots, b_n)$ 를 직교/직교 기저  $(u_1, \dots, u_n)$   $V$ . 이 기저는 항상 존재하고(Liesen and Mehrmann, 2015)  $\text{span}[b_1, \dots, b_n] = \text{스팬}[u_1, \dots, u_n]$ . 그람-슈미트 직교화 방법은 반복적으로 그람-슈미트  $V$ 의 임의의 기저  $(b_1, \dots, b_n)$ 로부터 직교 기저  $(u_1, \dots, u_n)$ 를 다음과 같이 구성합니다.

직교화

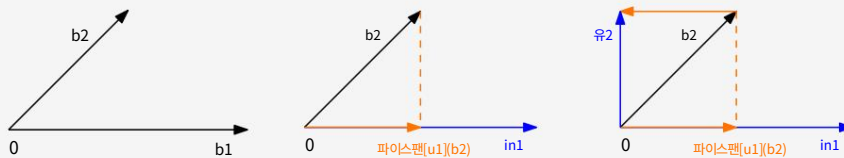
$$u_1 := b_1 \quad (3.67)$$

$$u_k := b_k - \pi_{\text{스팬}[u_1, \dots, u_{k-1}]}(b_k), \quad k = 2, \dots, n. \quad (3.68)$$

(3.68)에서  $k$ 번째 기저 벡터  $b_k$ 는 처음  $k-1$ 개의 구성된 직교 벡터  $u_1, \dots, u_{k-1}$ ; 섹션 3.8.2를 참조하십시오. 그런 다음 이 투영을  $b_k$ 에서 빼서  $u_1, \dots, u_{k-1}$ . 모든  $n$ 개의 기저 벡터  $b_1, \dots, b_n$ 에 대해 이 절차를 반복합니다.  $\dots, b_n$ 은  $V$ 의 직교 기저  $(u_1, \dots, u_n)$ 를 산출합니다.

$u_k$ 를 정규화하면  $k=1$ 에 대해  $\|u_k\| = 1$ 인 ONB를 얻습니다.  $\dots$

실시예 3.12(Gram-Schmidt 직교화)



(a) 원래의 비직교 (b) 첫 번째 새로운 기저 벡터 (c) 직교 기저 벡터  $u_1$  기저 벡터  $b_1, b_2$ .  $u_1 = b_1$  및  $b_2$  및  $u_2 = b_2 - \pi_{\text{스팬}[u_1]}(b_2)$ 의 투영.에 걸쳐 있는 부분 공간에

인.

$R^2$ 의 기저  $(b_1, b_2)$ 를 고려하십시오. , 어디

$$b_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad (3.69)$$

그림 3.12(a)도 참조하십시오. Gram-Schmidt 방법을 사용하여 다음과 같이  $R^2$ 의 직교 기저  $(u_1, u_2)$ 를 구성합니다(내적을 내적이라고 가정).

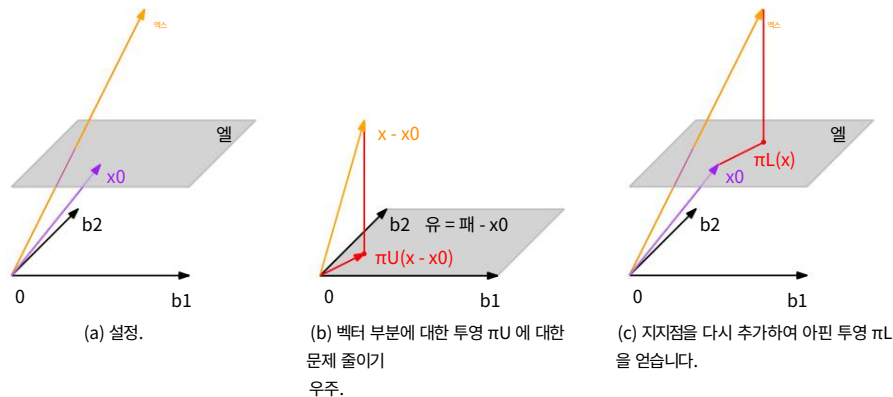
$$u_1 := b_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad (3.70)$$

$$u_2 := b_2 - \pi_{\text{스팬}[u_1]}(b_2) \stackrel{(3.45)}{=} b_2 - \frac{\langle u_1, b_2 \rangle}{\langle u_1, u_1 \rangle} u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1 \cdot 0 + 0 \cdot 1}{1^2 + 0^2} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (3.71)$$

그림 3.12 그람-슈미트 직교화. (a)  $R^2$ 의 비직교 기저  $(b_1, b_2)$ ; (b) 먼저 구성된 기저 벡터  $u_1$  및  $b_2$ 의 스패  $[u_1]$ 에 대한 직교 투영; (c)  $R^2$ 의 직교 기저  $(u_1, u_2)$ .

90

그림 3.13 아핀 공간으로의 투영. (a) 원래 설정; (b)  $x - x_0$  이 방향 공간  $U$ 에 투영될 수 있도록  $-x_0$  만큼 이동된 설정; (c) 투영은  $x_0 + \pi_U(x - x_0)$ 로 다시 변환되어 최종 직교 투영  $\pi_L(x)$ 를 제공합니다.



이러한 단계는 그림 3.12(b) 및 (c)에 설명되어 있습니다. 우리는  $u_1$  과  $u_2$  가 직교임을 즉시 알 수 있습니다. 즉,  $u_1^T u_2 = 0$ .

### 3.8.4 아핀 부분 공간에 투영 지금까지 우리는 벡터를 저

차원 부분 공간  $U$ 에 투영하는 방법에 대해 논의했습니다. 다음에서는 벡터를 아핀 부분 공간에 투영하는 솔루션을 제공합니다.

그림 3.13(a)의 설정을 고려하십시오. 아핀 공간  $L = x_0 + U$ 가 주어지며, 여기서  $b_1, b_2$  는  $U$ 의 기저 벡터입니다.  $x$ 의  $L$ 에 대한 직교 투영  $\pi_L(x)$ 를 결정하기 위해 문제를 해결 방법을 알고 있는 문제로 변환합니다. 벡터 부분 공간에 투영. 거기에 도달하기 위해 우리는  $x$ 와  $L$ 에서 지지점  $x_0$ 를 빼서  $L - x_0 = U$ 가 정확히 벡터 부분 공간  $U$ 가 되도록 합니다. 이제 섹션 3.8.2에서 논의한 부분 공간에 대한 직교 투영을 사용할 수 있습니다. 그림 3.13(b)에 나와 있는 투영  $\pi_U(x - x_0)$ 를 구합니다.

이 투영은 이제  $x_0$ 을 추가하여  $L$ 로 다시 변환될 수 있으므로 다음과 같이 아핀 공간  $L$ 에 대한 직교 투영을 얻습니다.

$$\pi_L(x) = x_0 + \pi_U(x - x_0), \quad (3.72)$$

여기서  $\pi_U(\cdot)$ 는 부분 공간  $U$ 에 대한 직교 투영, 즉  $L$ 의 방향 공간입니다. 그림 3.13(c) 참조.

그림 3.13에서 affine에서  $x$ 의 거리가 공간  $L$ 은  $U$ 에서  $x - x_0$ 의 거리와 동일합니다. 즉,

$$d(x, L) = \|x - \pi_L(x)\| = \|x - (x_0 + \pi_U(x - x_0))\| = \|x - x_0 - \pi_U(x - x_0)\| = \|\pi_U^\perp(x - x_0)\| \quad (3.73a)$$

$$d(x - x_0, \pi_U(x - x_0)) = d(x - x_0, U). \quad (3.73b)$$

섹션 12.1에서 분리 초평면의 개념을 유도하기 위해 아핀 부분 공간에 대한 투영을 사용할 것입니다.

## 3.9 회전

91

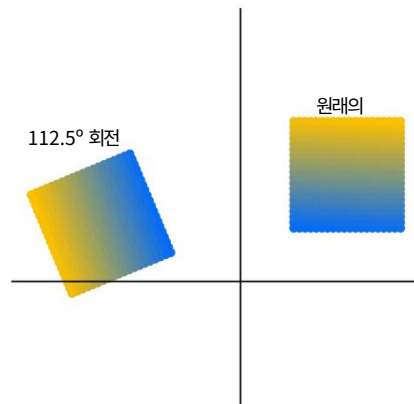


그림 3.14 회전 회전

원점에 대한 평면의 객체. 회전 각도가 양수이면 회전합니다.

시계 반대 방향.

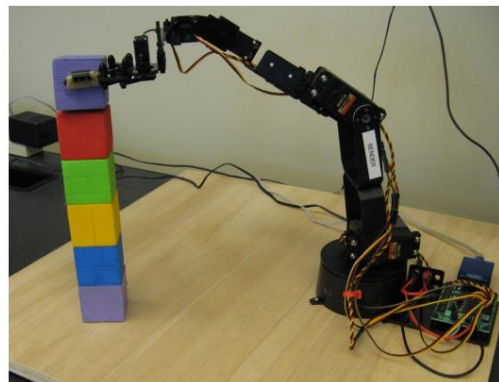


그림 3.15  
로봇팔이 필요하다  
물체를 집어 올리거나  
올바르게 배치하기  
위해 관절을 회전  
시킵니다.  
(Deisenroth  
et al., 2015)에서 가져온  
그림.

## 3.9 회전

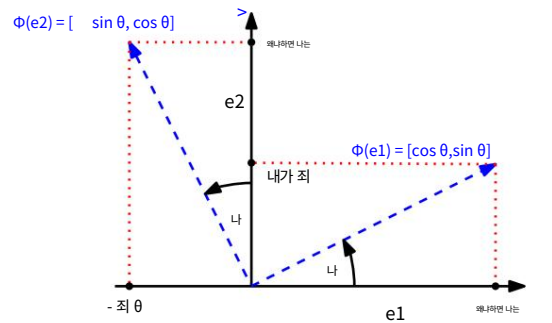
3.4절에서 논의된 길이와 각도 보존은 직교 변환 행렬을 사용한 선형 매핑의 두 가지 특성입니다. 다음에서는 회전을 설명하는 특정 직교 변환 행렬에 대해 자세히 살펴보겠습니다.

회전은 선형 매핑입니다(보다 구체적으로 회전의 자동형성 a Euclidean vector space) 원점에 대해 각도  $\theta$  만큼 평면을 회전시키는, 즉 원점이 고정된 점이다. 양의 각도  $\theta > 0$ 인 경우 일반적인 규칙에 따라 시계 반대 방향으로 회전합니다. 변환 행렬이 다음과 같은 예가 그림 3.14에 나와 있습니다.

$$R = \begin{pmatrix} 0.38 & 0.92 \\ 0.92 & 0.38 \end{pmatrix} \quad (3.74)$$

회전의 중요한 응용 분야에는 컴퓨터 그래픽과 로봇이 포함됩니다. 예를 들어, 로봇 공학에서는 물체를 집거나 배치하기 위해 로봇 팔의 관절을 회전하는 방법을 아는 것이 종종 중요합니다 (그림 3.15 참조).

그림 3.16 각도  $\theta$ 만큼  $R^2$ 의 표준 기저 회전.



### 3.9.1 $R^2$ 에서의 회전

표준 기저를 고려하십시오.  $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$   $R^2$ 의, 정의

$R^2$ 의 표준 좌표계 그림 3.16과 같이 이 좌표계를 각도  $\theta$ 만큼 회전시키는 것을 목표로 합니다. 회전된 벡터는 여전히 선형 독립이므로  $R^2$ 의 기저입니다. 이는 회전이 기저 변화를 수행함을 의미합니다.

회전 행렬

회전  $\Phi$ 는 선형 매핑이므로 회전 행렬  $R(\theta)$ 로 표현할 수 있습니다. 삼각법(그림 3.16 참조)을 사용하면  $R^2$ 의 표준 기저와 관련하여 회전된 축의 좌표( $\Phi$  이미지)를 결정할 수 있습니다. 우리는 얻는다

$$\Phi(e_1) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad \Phi(e_2) = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}. \quad (3.75)$$

따라서 회전 좌표  $R(\theta)$ 에 기저 변화를 수행하는 회전 행렬은 다음과 같이 주어집니다.

$$R(\theta) = \Phi(e_1) \quad \Phi(e_2) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (3.76)$$

### 3.9.2 $R^3$ 에서의 회전

$R^2$ 의 경우와 달리  $R^3$ 에서는 1차원 축을 중심으로 2차원 평면을 회전할 수 있습니다. 일반 회전 행렬을 지정하는 가장 쉬운 방법은 표준 기저  $e_1, e_2, e_3$ 의 이미지가 어떻게 회전되어야 하는지 지정하고 이러한 이미지  $Re_1, Re_2, Re_3$ 이 서로 직교하는지 확인하는 것입니다. 그런 다음 표준 기반의 이미지를 결합하여 일반 회전 행렬  $R$ 을 얻을 수 있습니다.

의미 있는 회전 각도를 가지려면 2차원 이상에서 작업할 때 "시계 반대 방향"이 무엇을 의미하는지 정의해야 합니다. 축에 대한 "반시계 방향"(평면) 회전은 "끝에서 원점을 향하여" 축을 볼 때 축에 대한 회전을 의미한다는 규칙을 사용합니다. 3개의 표준 기저 벡터에 대한  $R^3$  회전에서 (그림 3.17 참조):

, 따라서 세 가지(평면)가 있습니다.

## 3.9 회전

93

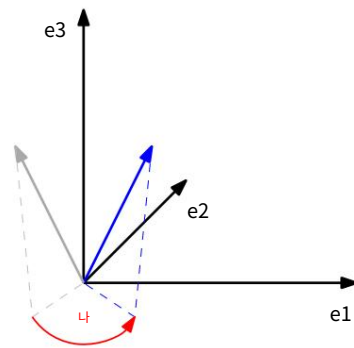


그림 3.17 e3축  
에 대한 각도  $\theta$ 만큼 R3의  
벡터(회색) 회전. 회전된 벡  
터는 다음과 같습니다.

푸른.

■ e1축에 대한 회전

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (3.77)$$

여기서, e1 좌표는 고정되어 있고, e2e3 평면에서 반시계 방향으로 회전이 수행된다.

■ e2축에 대한 회전

$$R_2(\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (3.78)$$

e2 축을 중심으로 e1e3 평면을 회전하는 경우 "끝"에서 원점을 향하는 e2 축을 봐야 합니다.

■ e3축에 대한 회전

$$R_3(\theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.79)$$

그림 3.17은 이것을 보여줍니다.

### 3.9.3 n 차원에서의 회전

2D 및 3D에서 n차원 Euclidean 벡터 공간으로의 회전 일반화는 n-2 차원을 고정하고 n차원 공간에서 2차원 평면으로 회전을 제한하는 것으로 직관적으로 설명할 수 있습니다. 3차원의 경우와 같이 모든 평면을 회전할 수 있습니다(Rn의 2차원 부분공간).

정의 3.11 (주어진 회전). V를 n차원 유클리드 벡터 공간이라고 하고  $\Phi: V \rightarrow V$  변환  $\Phi$ 를 갖는 자기형성



$$\text{행}(\theta) := \begin{pmatrix} 1 & 0 & 0 & \dots & \dots \\ 0 & \cos \theta & 0 & \sin \theta & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \sin \theta & 0 & 0 & \dots & \dots \\ 0 & \dots & \dots & 0 & \dots \end{pmatrix} \quad n \times n, \quad \text{인 } j$$

왜냐하면 나는

기븐스 로테이션

$$r_{ii} = \cos \theta, r_{ij} = -\sin \theta, r_{ji} = \sin \theta, r_{jj} = \cos \theta$$

### 3.9.4 회전의 속성

- 
- 
- 

### 3.10 추가 읽기

## 3.10 추가 읽기

95

커널 방법(Scholkopf 및 Smola는 많은 선형 알, 2002). 커널 방법은 고리증이 순전히 내부 제품 계산으로 표현될 수 있다는 사실을 알고 있습니다. 그런 다음 "커널 트릭"을 통해 이러한 (잠재적으로 무한 차원의) 피처에 암시적으로 내적 이 기능 공간을 명시적으로 알지도 못한 채 공간. 이것은 허용했다 다음과 같이 기계 학습에 사용되는 많은 알고리즘의 "비선형화" 차원 축소를 위한 kernel-PCA(Scholkopf et al. 1997). 가우스 시안 과정(Rasmussen and Williams, 2006)도 범주에 속합니다. 커널 방법의 최신 기술이며 확률적 회귀(데이터 포인트에 곡선 맞추기)의 최신 기술입니다. 커널의 아이디어를 탐구하다 12장에서 추가로 설명합니다.

투영은 예를 들어 음영 흐름을 생성하기 위해 컴퓨터 그래픽에서 자주 사용됩니다. 최적화에서 직교 투영은 종종 (반복적으로) 잔류 오류를 최소화합니다. 이것은 또한 기계 학습의 응용 프로그램이 있습니다. 예를 들어 선형 회귀에서 다음과 같은 (선형) 함수를 찾으려면 잔차 오류, 즉 선형 함수에 대한 데이터의 직교 투영 길이를 최소화합니다(Bishop, 2006). 우리는 이것을 9장에서 더 조사할 것입니다. PCA(Pearson, 1901; Hotelling, 1933) 또한 고차원 데이터의 차원을 줄이기 위해 투영을 사용합니다. 이에 대해서는 10장에서 더 자세히 논의할 것이다.

## 수업 과정

3.1 모든  $x = [x_1, x_2]$  에 대해 정의된  $h \cdot, \cdot i$  를 보여라.  $\cdot > \in R^2$  및  $y = [y_1, y_2] \cdot > \in R^2$  \_

$$hx, yi := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2(x_2 y_2)$$

내부 제곱입니다.

3.2  $R^2$  의 모든  $x$  와  $y$  에 대해 정의된  $h \cdot, \cdot i$  가 있는  $R^2$  를 다음 과 같이 고려하십시오.

$$y_i := x_1^2 - 2x_1 x_2 + x_2^2$$

$h \cdot, \cdot i$  는 내적인가?

3.3 사이의 거리를 계산

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, y = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

를 사

용하여  $hx, yi := x > y$

$$\text{비. } hx, yi := x > Ay, \quad A := \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

3.4 사이의 각도 계산

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \text{ 및 } y = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

를 사

용하여  $hx, yi := x$

$$> y \text{ b. } hx, yi := x > \text{ 기준, } B := \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$

3.5 내적이 있는 유클리드 벡터 공간  $R^5$  를 고려하십시오. 부분공간

$U \subseteq R^5$  및  $x \in R^5$  는

$$U = \text{스팬} \left[ \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ 5 \\ 0 \\ 7 \end{pmatrix}, \begin{pmatrix} -1 \\ -9 \\ 4 \\ 4 \\ 1 \end{pmatrix} \right], \quad x = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$$

$\perp$ .  $U$  b 에 대한  $x$  의 직교 투영  $\pi_U(x)$  를 결정 합니다. 거리  $d(x, U)$  결  
정

3.6 내적과 함께  $R^3$  를 고려하라

$$hx, yi := x \cdot \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

또한  $e_1, e_2, e_3$  을  $R^3$  의 표준/정규 기반으로 정의 합니다.

†.  $e_2$ 의 직교 투영  $\pi_U(e_2)$ 를 결정합니다.

$$U = \text{스팬}\{e_1, e_3\}.$$

힌트: 직교성은 내적을 통해 정의됩니다. 비. 거리  $d(e_2, U)$ 를 계산합니다. 씨. 시나리오 그리기: 표준 기저 벡터 및  $\pi_U(e_2)$

3.7  $V$ 를 벡터 공간이라고 하고  $\pi$ 를  $V$ 의 엔도모피즘이라고 하자.

†.  $\text{id}_V - \pi$ 가 투영인 경우에만  $\pi$ 가 투영임을 증명합니다. 여기서  $\text{id}_V$ 는  $V$ 에 대한 동일성 내형성입니다.

비. 이제  $\pi$ 가 투영이라고 가정합니다.  $\text{Im}(\text{id}_V - \pi)$  및  $\ker(\text{id}_V - \pi)$ 를  $\text{Im}(\pi)$  및  $\ker(\pi)$ 의 함수로 계산합니다.

3.8 Gram-Schmidt 방법을 사용하여 2차원 부분공간  $U \subseteq \mathbb{R}^3$ 의 기저  $B = (b_1, b_2)$ 를  $U$ 의 ONB  $C = (c_1, c_2)$ 로 변환합니다. 여기서

$$b_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad b_2 := \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

3.9  $n \in \mathbb{N}$ 이라고 하고  $x_1, \dots, x_n > 0$ 은  $x_1 + \dots + x_n = 1$ 이 되도록  $n$ 개의 양의 실수입니다. Cauchy-Schwarz 부등식을 사용하여 다음을 보여줍니다.

$$\text{가. } \sum_{i=1}^n x_i \geq \frac{1}{n} \\ \text{비. } \sum_{i=1}^n \frac{1}{x_i} \geq n$$

트:  $\mathbb{R}^n$ 의 내적을 생각해 보세요. 그런 다음 특정 벡터  $x, y \in \mathbb{R}^n$ 을 선택하고 Cauchy-Schwarz 부등식을 적용합니다.

3.10 벡터 회전

$$x_1 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad x_2 := \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

30°.