

---

## 내용물

머리말	1
1부 수학 기초	9
소개 및 동기	11
1 1.1 직관을 위한 단어 찾기 1.2 이 책을 읽는	12
두 가지 방법 1.3 연습 문제 및 피드백	13
	16
방정식 시스템과 행렬	17
행 방정식 풀기 시스템 2.4 벡터 공간 2.5 선형	19
독립 2.6 기저 및 순위 2.7 선형 매핑 2.8 아핀	22
공간 2.9 추가 읽기 연습	27
	35
	40
	44
	48
	61
	63
	64
삼     해석 기하학 3.1 규범 3.2	70
내적 3.3 길이와 거리 3.4 각도와 직	71
교 3.5 직교 기저 3.6 직교 보수 3.7	72
함수의 내적 3.8 직교 투영 3.9 회전	75
3.10 추가 읽기 연습	76
	78
	79
	80
	81
	91
	94
	96
4 행렬 분해 4.1 행렬식과 자취	98
	99

4.2 고유값과 고유 벡터 4.3 쉐레스키 분해 4.4 고유 분해와 대각화 4.5 특이값 분해 4.6 행렬 근사 4.7 행렬 계통 발생 4.8 추가 읽기 연습	105 114 115 119 129 134 135 137
변량 함수의 다변량 정규 분포 5.1 다변량 정규 분포의 기울기 5.2 벡터 값 함수의 기울기 5.3 행렬의 기울기 5.4 기울기 계산에 유용한 항등식 5.5 역전파 및 자동 미분 5.6 더 높은 선형 및 5차 선형 미분 테일러 읽기	139 141 146 149 155 158 159 164 165 170 170
6 확률과 분포 6.1 확률 공간의 구성 6.2 이산 및 연속 확률 6.3 합법칙, 곱법칙, 베이즈 정리 6.4 요약 통계 및 독립성 6.5 가우스 분포 6.6 켈레와 지수군 6.7 변수의 변화/역 읽기 변환 6.8 수업 과정	172 172 178 183 186 197 205 214 221 222
한 최적화 연속 최적화 조건 최적화 방법 최적화 알고리즘 7.3 블록 최적화 7.4 추가 읽기 연습	225 227 233 236 246 247
<b>2부 중앙 머신 러닝 문제</b>	<b>249</b>
8 모델이 데이터를 만날 때 8.1 데이터, 모델 및 학습 8.2 경험적 위험 최소화 8.3 매개변수 추정 8.4 확률적 모델링 및 추론 8.5 방향성 그래픽 모델	251 251 258 265 272 278

내용물	iii
8.6 모델 선택	283
9 선형 회귀 9.1 문제 공식화 9.2 매	289
개변수 추정 9.3 베이지안 선형 회귀	291
9.4 직교 투영으로서의 최대 가능성	292
9.5 추가 읽기	303
	313
	315
10 주성분 분석을 통한 차원 축소 317	
10.1 문제 설정 318	
10.2 최대 분산 관점 320	
10.3 투사 관점 325	
10.4 고유벡터 계산과 낮은 순위 근사치 333	
10.5 고차원의 PCA 335	
10.6 실제 PCA의 주요 단계 336	
10.7 잠재 변수 관점 339	
10.8 더 읽을거리 343	
11 가우스 혼합 모델을 사용한 밀도 추정 11.1 가우시안 혼합 모델 11.2 최대 가	348
능성을 통한 매개변수 학습 11.3 EM 알고리즘 11.4 잠재 변수 관점 11.5 추가 읽	349
기	350
	360
	363
	368
12 서포트 벡터 머신을 사용한 분류 12.1 초평면 분리 12.2 원시 서포트	370
벡터 머신 12.3 이중 서포트 벡터 머신 12.4 커널 12.5 수치 해 12.6 추	372
가 자료	374
	383
	388
	390
	392
참고문헌	395



---

## 머리말

머신 러닝은 인간을 정제하려는 오랜 시도 중 가장 최근의 것입니다. 지식과 추론을 기계를 구성하고 자동화 시스템을 엔지니어링하는 데 적합한 형태로 변환합니다. 머신러닝이 되면서 보편화되고 소프트웨어 패키지가 사용하기 쉬워짐에 따라 낮은 수준의 기술적 세부 사항을 추상화하는 것이 자연스럽고 바람직합니다. 그리고 실무자에게 숨겨져 있습니다. 그러나 이것은 위험을 수반합니다. 실무자가 설계 결정을 알지 못하므로, 머신러닝 알고리즘의 한계 더 많은 것을 배우고자 하는 열정적인 실무자 성공적인 기계 학습 알고리즘 뒤에 숨겨진 마법은 현재 직면하고 있습니다. 필수 필수 지식의 어려운 세트:

- 프로그래밍 언어 및 데이터 분석 도구
- 대규모 계산 및 관련 프레임워크
- 수학 및 통계, 그리고 이를 기반으로 하는 기계 학습 방법

대학에서는 기계 학습에 대한 입문 과정에 지출하는 경향이 있습니다. 이러한 전제 조건 중 일부를 다루는 과정의 초기 부분. 역사적 이유로 기계 학습 과정은 컴퓨터에서 가르치는 경향이 있습니다. 학생들이 종종 처음 두 영역에서 훈련을 받는 과학부 지식에 대해서는 그렇지 못하지만 수학이나 통계에는 그다지 많지 않습니다. 현재 기계 학습 교과서는 주로 기계 학습 알고리즘과 방법론에 중점을 두고 있으며 독자가 수학 및 통계에 능숙하다고 가정합니다. 따라서 이 책들은

처음에 배경 수학에 대한 하나 또는 두 개의 장 책의 또는 부록으로. 우리는 원하는 많은 사람들을 찾았습니다. 기계 학습을 읽는 데 필요한 수학적 지식과 씨름하는 기본적인 기계 학습 방법의 기초를 탐구합니다. 교과서. 대학에서 학부 및 대학원 과정을 가르쳤던 우리는 고등학교 수학과 표준 기계 학습 교과서를 읽는 데 필요한 수학 수학 수준의 격차가 너무 크다는 것을 발견했습니다.

많은 사람들에게 큰.

이 책은 기본적인 머신러닝 개념의 수학적 기초를 전면에 내세우고 정보를 한곳에 모아 서 이 기술 격차가 좁혀지거나 심지어 닫힙니다.

## 왜 기계 학습에 대한 또 다른 책인가?

기계 학습은 직관적으로 명백해 보이지만 공식화하기는 놀라울 정도로 어려운 개념을 표현하기 위해 수학 언어를 기반으로 합니다. 제대로 공식화되면 해결하려는 작업에 대한 통찰력을 얻을 수 있습니다. 전 세계 수학 학생들의 공통된 불만 중 하나는 다루는 주제가 실제 문제와 거의 관련이 없는 것 같다는 것입니다. 우리는 기계 학습이 사람들이 수학을 배우도록 하는 명백하고 직접적인 동기라고 믿습니다.

“수학은 대중의 마음에서 공포증과 불안과 연결되어 있습니다. 당신은 우리가 거미에 대해 이야기하고 있다고 생각할 것입니다.” (Strogatz, 2014, 281페이지)

이 책은 현대 머신 러닝의 기초를 형성하는 방대한 수학 문헌에 대한 가이드북이 되기 위한 것입니다. 우리는 근본적인 기계 학습 문제의 맥락에서 수학 개념의 유용성을 직접 지적함으로써 수학적 개념의 필요성을 자극합니다. 책을 짧게 유지하기 위해 많은 세부 사항과 고급 개념이 생략되었습니다. 여기에 제시된 기본 개념과 기계 학습의 더 큰 맥락에 어떻게 부합하는지 독자는 각 장의 끝에서 제공하는 추가 연구를 위한 수많은 리소스를 찾을 수 있습니다. 수학적 배경이 있는 독자를 위해 이 책은 기계 학습에 대해 간략하지만 정확하게 설명되어 있습니다. 기계 학습의 방법과 모델에 초점을 맞춘 다른 책들과 대조적으로(MacKay, 2003; Bishop, 2006; Alpaydin, 2010; Barber, 2012; Murphy, 2012; Shalev-Shwartz와 Ben-David, 2014; Rogers와 Girolami, 2016) 또는 기계 학습의 프로그래밍 방식(Muller 및 Guido, 2016; Raschka 및 Mirjalili, 2017; Chollet 및 Allaire, 2018)에서 기계 학습 알고리즘의 대표적인 예는 4개만 제공합니다. 대신 우리는 모델 자체의 수학적 개념에 초점을 맞춥니다. 독자들이 기계 학습의 기본 질문에 대해 더 깊이 이해하고 기계 학습 사용에서 발생하는 실용적인 질문을 수학적 모델의 기본 선택과 연결할 수 있기를 바랍니다.

우리는 고전적인 기계 학습 책을 쓰는 것을 목표로 하지 않습니다. 대신, 우리의 의도는 4가지 핵심 머신 러닝 문제에 적용되는 수학적 배경을 제공하여 다른 머신 러닝 교과서를 더 쉽게 읽을 수 있도록 하는 것입니다.

## 대상 고객은 누구입니까?

기계 학습의 적용이 사회에 널리 보급됨에 따라 모든 사람이 기계 학습의 기본 원리를 어느 정도 이해해야 한다고 믿습니다. 이 책은 학문적 수학적 스타일로 작성되어 기계 학습의 이면에 있는 개념에 대해 정확히 알 수 있습니다. 우리는 이 간결해 보이는 스타일에 익숙하지 않은 독자들이 인내심을 갖고 각 주제의 목표를 염두에 두도록 권장합니다. 큰 그림과 관련하여 유용한 지침을 제공하기를 희망하여 텍스트 전체에 주석과 주석을 뿌립니다.

이 책은 독자가 일반적으로 수학적 지식을 가지고 있다고 가정합니다.

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.

머리말

삼

고등학교 수학과 물리학을 다룬다. 예를 들어 독자는 이전에 도함수와 적분, 기하 벡터를 보았어야 합니다. 2차원 또는 3차원에서. 거기에서 시작하여 이러한 개념을 일반화합니다. 따라서 이 책의 대상 독자는 학부생을 포함합니다. 대학생, 야간학습자 및 온라인으로 참여하는 학습자 기계 학습 과정.

음악에 비유하자면 사람들이 하는 상호작용에는 세 가지 유형이 있습니다. 기계 학습과 함께:

Astute Listener 오픈 소스 소프트웨어, 온라인 자습서 및 클라우드 기반 도구의 제공으로 머신 러닝이 민주화되어 사용자는 파이프라인의 세부 사항에 대해 걱정할 필요가 없습니다. 사용자가 집중할 수 있는 기성 도구를 사용하여 데이터에서 통찰력을 추출합니다. 이를 통해 기술에 정통하지 않은 도메인 전문가가 기계 학습의 이점을 누릴 수 있습니다. 이것은 음악을 듣는 것과 유사합니다. 사용자는 다음 중에서 선택하고 식별할 수 있습니다. 다양한 유형의 머신 러닝과 이점을 제공합니다. 경험이 많은 사용자는 음악 평론가와 같으며 개인의 윤리, 공정성, 프라이버시 등 머신러닝의 사회 적용 이 책이 다음의 기초가 되기를 바랍니다.

머신 러닝의 인증 및 위험 관리에 대한 생각  
시스템을 구축하고 도메인 전문 지식을 사용하여 더 나은 기계 학습 시스템.

숙련된 아티스트 기계 학습의 숙련된 실무자는 플러그를 꽂을 수 있습니다. 다양한 도구와 라이브러리를 분석 파이프라인으로 재생합니다. 스테레오 일반 실무자는 다음을 이해하는 데이터 과학자 또는 엔지니어입니다. 머신 러닝 인터페이스와 그 사용 사례를 수행할 수 있습니다. 데이터에서 예측의 놀라운 위업. 이것은 고도로 숙련된 연주자가 기존 악기에 생명을 불어넣고 청중에게 즐거움을 선사할 수 있는 음악을 연주하는 것과 유사합니다. 여기에 제시된 수학을 입문서로 사용하여 실무자는 자신이 선호하는 방법의 이점과 한계를 이해하고 확장 및 확장할 수 있습니다.

기존 기계 학습 알고리즘을 일반화합니다. 우리는 이 책이 보다 엄격하고 원칙적인 개발을 위한 추진력을 제공합니다. 기계 학습 방법.

신생 작곡가 머신러닝이 새로운 영역에 적용되면서, 기계 학습 개발자는 새로운 방법을 개발하고 확장해야 합니다. 기존 알고리즘. 그들은 종종 이해가 필요한 연구원입니다. 기계 학습의 수학적 기초와 서로 다른 작업 간의 관계를 밝혀냅니다. 이것은 음악 작곡가와 유사합니다. 음악 이론의 규칙과 구조, 새롭고 놀라운 작품을 만듭니다. 이 책이 다른 기술 서적에 대한 높은 수준의 개요를 제공하기를 바랍니다. 머신 러닝의 작곡가가 되고 싶은 사람들을 위해. 있다 제안하고 제안할 수 있는 새로운 연구원에 대한 사회의 큰 필요 학습의 많은 도전 과제를 공격하기 위한 새로운 접근 방식 탐색 데이터에서.

## Acknowledgements 우리는

책의 초기 초안을 보고 고통스러운 개념 설명으로 고통을 겪은 많은 사람들에게 감사합니다. 우리는 격렬하게 동의하지 않는 그들의 아이디어를 구현하려고 노력했습니다. 특히 이 책의 많은 부분을 주의 깊게 읽고 구조와 표현에 대한 상세한 제안을 한 Christfried Webers에게 감사드립니다. 많은 친구와 동료들도 각 장의 다른 버전에 시간과 에너지를 제공할 만큼 친절했습니다.

<https://github.com>을 통해 개선 사항을 제안한 온라인 커뮤니티의 관대함 덕분에 운이 좋았습니다. 책을 크게 개선했습니다.

다음 사람들은 <https://github.com> 을 통해 버그를 발견하고 설명을 제안하고 관련 문헌을 제안했습니다. 또는 개인 통신. 그들의 이름은 알파벳순으로 정렬되어 있습니다.

Abdul-Ganiy Usman  
Adam Gaier Adele  
Jackson Aditya  
Menon Alasdair Tran  
Aleksandar Krnjaic  
Alexander  
Makrigiorgos Alfredo  
Canziani Ali Shafiti Amr  
Khalifa Andrew Tanggara  
Angus Gruen Antal A.  
Buss Antoine Toisoul Le  
Cann Areg Sarvazyan  
Artem Artemev

아르티움 스테파노프  
빌 크로마다스  
밥 윌리엄슨  
분핑 임  
차오 꾸  
첵 리  
크리스 설록  
크리스토퍼 그레이  
다니엘 맥나마라  
다니엘 우드  
대런 시겔  
데이비드 존스턴  
첸 다웨이

엘렌 브로드  
Fengkuangtian Zhu  
피오나 콘돈  
게오르기오스 테오도루  
허신  
아이린 라이사 카메니  
야콥 나바글로  
제임스 헨스만  
제이미 리우  
존 카두어  
장 폴 에베제르  
제리 창  
자테쉬 신다레  
존 로이드  
조나스 응나웨  
존 마틴  
저스틴 시  
카이 아를쿠마란  
카밀 드레츠코프스키  
릴리 왕  
라이오넬 톤지 응우페유  
리디아 크누핑 “  
마흐무드 아슬란  
마크 하텐슈타인  
마크 반 데어 빌크  
마커스 헤글란트  
마틴 휴잉  
매튜 알저  
매튜 리



머리말

5

막시무스 맥캔

장명옌

마이클 베넷

마이클 페더슨

신민정

모하마드 말렉자데

나빈 쿠마르

니코 몬탈리

오스카 팔

패트릭 헨릭센

패트릭 비스콜렉

파타라왓 초마이

폴 켈리

페트로스 크리스토둘루

표트르 아누체프스키

프라나프 수브라마니

꾸유콩

라깅 시간

루이 장

라이언-리스 그리피스

솔로몬 카봉고

사무엘 오군몰라

산딕 마바디아

사르베시 니콤브

세바스찬 라슈카

세나나약 세시 쿠마르 카리

백승헌

샤바즈 차우드하리

샤키르 모하메드

손 베리

셰이크 압둘 라힘 알리

쉬에 성

스리다르 티아가라잔

시에드 누만 하산

시몬 브리치

토마스 볼러 "

티무르 샤라포프

툼 멜라메드

빈센트 아담

Vincent Dutordoir

부민

와심 아프타브

웬지

보이치에흐 스토크비에츠

총 샤오난

장 샤오웨이

야저우 하오

루오 이청

이영

유 루

윤청

황위샤오

잭 크랭코

차오 자젠

조 놀란

GitHub를 통해 실명이 나열되지 않은 기고자

GitHub 프로파일은 다음과 같습니다.

SamDataMad

bumpitiousmonkey

idoamihai deepakiim

인싸

HorizonP

cs-maillist

kudo23

엠펏 빅터

비간드

17SKYE 제

장1995

우리는 또한 Parameswaran Raman과 Cambridge University Press에서 조직한 많은 익명의 검토자들에게 매우 감사합니다. 이들은 원고의 이전 버전에서 하나 이상의 장을 읽고 상당한 개선으로 이어지는 건설적인 비평을 제공했습니다. LATEX 관련 문제 에 대한 상세하고 신속한 조언을 위해 LATEX 지원팀인 Dinesh Singh Negi에게 특별 언급이 있습니다. 마지막으로 이 책의 잉태 과정을 참을성 있게 안내해 주신 편집자 Lauren Cowles에게 매우 감사드립니다.

## 기호 표

기호 a, b,	전형적인 의미
c, b, c	스칼라는 소문자입니다
x, y, z	벡터는 굵은 소문자입니다.
A, B, C	행렬은 굵은 대문자입니다.
$\mathbf{x}^T, \mathbf{y}^T$	벡터 또는 행렬의 전치
$\mathbf{x} \cdot \mathbf{y}$	역행렬
$\langle \mathbf{x}, \mathbf{y} \rangle$	x 와 y 의 내적
$\mathbf{x} \cdot \mathbf{y}$	x 와 y 의 내적
$\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ (순서된) 튜플	
$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$ 가로로 쌓인 열 벡터의 행렬	
$\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ 벡터 세트(순서 없음)	
지, 엔	각각 정수와 자연수
R, C	실수 및 복소수, 각각
$\mathbb{R}^n$	실수의 n차원 벡터 공간
$\forall x$	범용 수량자: 모든 x 에 대해
$\exists x$	실존 수량자: 존재 x
$x \in b$	b 로 정의 된다
$a \sim b$	b 는 다음과 같이 정의됩니다 .
$a \propto b$	a 는 b에 비례합니다 . 즉, $a = \text{상수} \cdot b$
$g \circ f$	기능 구성: "g 다음에 f"
$\Leftrightarrow$	만약 그리고 만약에
$\Rightarrow$	암시
A, C	세트
$a \in A$	a 는 집합 A 의 요소입니다.
$\emptyset$	빈 세트
$A \setminus B$	B가 없는 A: A에는 있지만 B에는 없는 요소 집합
d	차원 수; $d = 1, \dots, d$
N	데이터 포인트 수 $n = 1, \dots, N$
나는	크기가 $m \times m$ 인 단위 행렬
$\mathbf{0}_{m,n}$	크기가 $m \times n$ 인 0으로 구성된 행렬
$\mathbf{1}_{m,n}$	$m \times n$ 크기의 행렬
그의 희	표준/정규 벡터(여기서 i 는 1인 구성요소)
미한	벡터 공간의 차원
$\text{rk}(A)$	행렬 A 의 순위
$\text{Im}(\Phi)$	선형 매핑 $\Phi$ 의 이미지
$\text{ker}(\Phi)$	선형 매핑 $\Phi$ 의 커널(null space)
스팬[b1]	b1 의 스패(생성 집합)
$\text{tr}(A)$	A의 흔적
$\det(A)$	A의 행렬식
$\ \cdot\ $	절대값 또는 결정자(컨텍스트에 따라 다름)
$\ \cdot\ _2$	표준; 지정하지 않는 한 유클리드
$\ \cdot\ _1$	고유값 또는 라그랑주 승수
$\ \cdot\ _\lambda$	고유값 $\lambda$ 에 해당하는 고유 공간

머리말

7

기호 $x \perp$	전형적인 의미
$y$	벡터 $x$ 와 $y$ 는 직교합니다.
$\mathcal{V}$	벡터 공간
$\mathcal{V}^\perp$	벡터 공간 $\mathcal{V}$ 의 직교 보수
$\sum_{n=1}^N x_n$	$x_n$ 의 합 : $x_1 + \dots + x_N$
$\sum_{n=1}^N x_n$	$x_n$ 의 곱 : $x_1 \cdot \dots \cdot x_N$
$\frac{\partial f}{\partial x}$	$x$ 에 대한 $f$ 의 편도 함수
$\frac{df}{dx}$	$x$ 에 대한 $f$ 의 총 도함수
$\nabla$ 그래디언트	
$f^* = \text{최소값 } f(x)$	$f$ 의 가장 작은 함수 값
$x^* \in \arg \min_x f(x)$	$f$ 를 최소화하는 $x^*$ 값 (참고: $\arg \min$ 은 값 집합을 반환함)
$L$	라그랑지안
$L$ 음의 로그 가능성	
$\sum_{k=1}^n$	이항 계수, $n$ k 선택
$VX[x]$	확률 변수 $X$ 에 대한 $x$ 의 분산
$EX[x]$	확률 변수 $X$ 에 대한 $x$ 의 기대치
$\text{Cov}X, Y [x, y]$	$x$ 와 $y$ 사이의 공분산 .
$X \perp\!\!\!\perp Y   Z$	$X$ 는 $Z$ 가 주어지면 $Y$ 와 조건부로 독립적입니다.
$X \sim \pi$	확률 변수 $X$ 는 $\pi$ 에 따라 분포됩니다.
$N(\mu, \Sigma)$	평균 $\mu$ 와 공분산 $\Sigma$ 를 갖는 가우스 분포
베르( $\mu$ )	모수 $\mu$ 가 있는 베르누이 분포
$\text{Bin}(N, \mu)$	매개변수 $N, \mu$ 가 있는 이항 분포
베타( $a, b$ )	매개변수 $a, b$ 가 있는 베타 분포

## 약어 및 두문자어 표

약어 의미	
예	예를 들어 (라틴어: 예를 들어)
GMM	가우스 혼합 모델
즉	Id est (라틴어: 이것은 의미)
이드	독립적, 동일하게 분포
사후 최대 MAP	
MLE	최대 가능도 추정/추정기
온앤비	직교 기저
PCA	주요 구성 요소 분석
PPCA	확률적 주성분 분석
참조	행 계층 형식
SPD	대칭, 양의 정부호
SVM	서포트 벡터 머신



## 1부

---

### 수학 기초



# 1

---

## 소개 및 동기

머신 러닝은 데이터에서 가치 있는 정보를 자동으로 추출하는 알고리즘을 설계하는 것입니다. 여기서 강조점은 "자동"에 있습니다. 즉, 기계 학습은 의미 있는 무언가를 생성하면서 많은 데이터 세트에 적용될 수 있는 범용 방법론에 관심이 있습니다. 머신 러닝의 핵심에는 데이터, 모델, 학습이라는 세 가지 개념이 있습니다.

머신 러닝은 본질적으로 데이터 기반이므로 데이터가 핵심 데이터에 있습니다. 기계 학습의. 머신 러닝의 목표는 데이터에서 가치 있는 패턴을 추출하는 범용 방법론을 설계하는 것입니다. 예를 들어, 방대한 양의 문서(예: 많은 도서관의 책)가 있는 경우 기계 학습 방법을 사용하여 문서 간에 공유되는 관련 주제를 자동으로 찾을 수 있습니다(Hoffman et al., 2010). 이 목표를 달성하기 위해 우리는 주어진 데이터 세트를 모델링 하는 것과 유사하게 일반적으로 데이터를 생성하는 프로세스와 관련된 모델을 설계합니다. 예를 들어 회귀 설정에서 모델은 입력을 실제 값 출력에 매핑하는 함수를 설명합니다. Mitchell(1997)의 말을 바꾸면: 데이터를 고려한 후 주어진 작업에 대한 성능이 향상되면 모델은 데이터에서 학습한다고 합니다.

목표는 우리가 미래에 관심을 가질 수 있는 아직 보이지 않는 데이터로 잘 일반화되는 좋은 모델을 찾는 것입니다. 학습은 모델의 매개변수를 최적화하여 데이터에서 패턴과 구조를 자동으로 찾는 학습 방법으로 이해할 수 있습니다.

기계 학습은 많은 성공 사례를 보았고 풍부하고 유연한 기계 학습 시스템을 설계하고 훈련하는 데 소프트웨어를 쉽게 사용할 수 있지만, 우리는 기계 학습의 수학적 기초가 더 복잡한 기계 학습 시스템의 기본 원리를 이해하는 데 중요하다고 믿습니다. 구축됩니다. 이러한 원칙을 이해하면 새로운 기계 학습 솔루션을 만들고, 기존 접근 방식을 이해하고 디버깅하고, 우리가 작업 중인 방법론의 고유한 가정과 한계에 대해 학습하는 데 도움이 될 수 있습니다.

## 1.1 직관을 위한 단어 찾기

기계 학습에서 우리가 정기적으로 직면하는 문제는 개념과

단어는 미끄럽고 기계 학습의 특정 구성 요소

시스템은 다른 수학적 개념으로 추상화될 수 있습니다. 예를 들어,

"알고리즘"이라는 단어는 기계 학습의 맥락에서 적어도 두 가지 다른 의미로 사용됩니다.

첫 번째 의미에서 우리는 "기계

학습 알고리즘"은 입력된 데이터를 기반으로 예측을 수행하는 시스템을 의미합니다. 이러한

알고리즘을 예측 변수라고 합니다. 두 번째 의미에서,

우리는 "머신 러닝 알고리즘"과 똑같은 문구를 사용합니다.

예측 변수의 일부 내부 매개변수를 조정하는 시스템

미래의 보이지 않는 입력 데이터에 대해 잘 수행합니다. 여기서 우리는 이 적응을 시스템 훈련이라고 합니다.

이 책은 모호성 문제를 해결하지 못하지만 문맥에 따라 동일한 표현이 가능하다는 점을 강조하고 싶습니다.

다른 것을 의미합니다. 그러나 우리는 모호함의 수준을 줄이기 위해 문맥을 충분히 명확하게 하려고 합니다.

이 책의 첫 부분은 수학 개념을 소개하고

기계의 세 가지 주요 구성 요소에 대해 이야기하는 데 필요한 기초

학습 시스템: 데이터, 모델 및 학습. 우리는 이것을 간략하게 설명 할 것입니다

여기에서 구성 요소를 살펴보고 8장에서 다시 살펴보겠습니다.

필요한 수학적 개념에 대해 논의했습니다.

모든 데이터가 수치는 아니지만 데이터를 다음과 같이 고려하는 것이 종종 유용합니다.

숫자 형식. 이 책에서는 데이터가 이미

컴퓨터 프로그램으로 읽기에 적합한 수치 표현으로 적절하게 변환됩니다. 따라서 우리는 데이터를 벡터로 생각합니다.처럼

단어가 얼마나 미묘한지 보여주는 또 다른 예는 (적어도) 세 가지가 있습니다.

벡터에 대해 생각하는 다양한 방법: 벡터를 숫자 배열(a

컴퓨터 과학 관점), 방향과 크기가 있는 화살표로서의 벡터(물리학적 관점), 덧셈과 덧셈을 따르는 객체로서의 벡터

스케일링(수학적 관점).

모델은 일반적으로 현재 데이터 세트와 유사한 데이터 생성 프로세스를 설명하는 데 사용됩니다. 따라서 좋은 모델도 생각할 수 있습니다.

실제(알 수 없는) 데이터 생성 프로세스의 단순화된 버전으로,

데이터 모델링 및 추출과 관련된 측면 캡처

그 안에 숨겨진 패턴. 그러면 좋은 모델을 사용하여 다음을 예측할 수 있습니다.

실제 경험을 수행하지 않고 현실 세계에서 일어날 것입니다.

멘션.

이제 문제의 핵심인 학습 구성 요소에 도달했습니다.

기계 학습. 데이터 세트와 적절한 모델이 주어진다고 가정합니다.

모델을 훈련 시킨다는 것은 효용 함수와 관련하여 모델의 일부 매개변수를 최적화하는 데 사용할 수 있는 데이터를 사용하는 것을 의미합니다.

모델은 훈련 데이터를 예측합니다. 대부분의 훈련 방법은

정상에 도달하기 위해 언덕을 오르는 것과 유사한 접근 방식으로 생각됩니다.

이 비유에서 언덕의 정점은 일부의 최대값에 해당합니다.



원하는 성능 측정. 그러나 실제로 우리는 보이지 않는 데이터에서 잘 수행되는 모델에 관심이 있습니다. 이미 본 데이터(훈련 데이터)에 대해 잘 수행한다는 것은 데이터를 암기하는 좋은 방법을 찾았음을 의미할 수 있습니다. 그러나 이것은 보이지 않는 데이터에 잘 일반화되지 않을 수 있으며 실제 적용에서 우리는 종종 기계 학습 시스템을 이전에 접하지 못한 상황에 노출해야 합니다.

이 책에서 다루는 기계 학습의 주요 개념을 요약해 보겠습니다.

- 데이터를 벡터로 표현합니다.
- 우리는 확률 또는 최적화 보기를 사용하여 적절한 모델을 선택합니다.
- 모델이 훈련에 사용되지 않은 데이터에 대해 잘 수행된다는 목표로 수치 최적화 방법을 사용하여 사용 가능한 데이터에서 학습합니다.

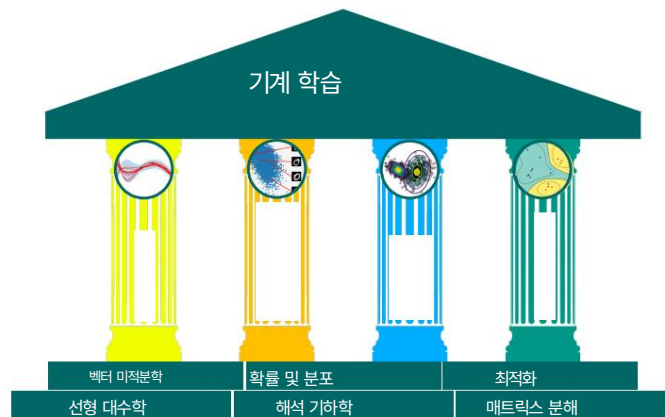
### 1.2 이 책을 읽는 두 가지 방법 우리는 기계

학습을 위한 수학을 이해하기 위한 두 가지 전략을 고려할 수 있습니다.

- **상향식:** 기본 개념부터 고급 개념까지 구축합니다. 이것은 종종 수학과 같은 보다 기술적인 분야에서 선호되는 접근 방식입니다. 이 전략은 독자가 항상 이전에 배운 개념에 의존할 수 있다는 장점이 있습니다. 불행히도, 실무자에게 많은 기본 개념은 그 자체로는 특별히 흥미롭지 않으며 동기가 부족하다는 것은 대부분의 기본 정의를 빨리 잊어버린다는 것을 의미합니다.
- **하향식:** 실제 요구 사항에서 보다 기본적인 요구 사항으로 드릴다운합니다. 이 목표 중심 접근 방식은 독자가 특정 개념에 대해 작업해야 하는 이유를 항상 알고 있고 필요한 지식의 경로가 명확하다는 장점이 있습니다. 이 전략의 단점은 지식이 잠재적으로 불안정한 기반 위에 구축되고 독자가 이해할 수 없는 단어 집합을 기억해야 한다는 것입니다.

우리는 이 책을 두 가지 방법으로 읽을 수 있도록 응용 프로그램에서 기초(수학적) 개념을 분리하는 모듈 방식으로 이 책을 작성하기로 결정했습니다. 이 책은 두 부분으로 나뉩니다. 1부는 수학적 기초를 다루고 2부는 1부의 개념을 그림 1.1과 같이 기계 학습의 4가지 기둥인 회귀, 차원을 형성하는 기본적인 기계 학습 문제 세트에 적용합니다. 감소, 밀도 추정 및 분류. 1부의 챕터는 대부분 이전 챕터를 기반으로 하지만 필요한 경우 챕터를 건너뛰고 뒤로 작업할 수 있습니다. 파트 II의 장은 느슨하게 연결되어 있고 어떤 순서로든 읽을 수 있습니다. 많은 포인터가 앞뒤로 있습니다.

그림 1.1  
기초 및  
네 개의 기둥  
기계 학습.



수학 개념을 연결하기 위해 책의 두 부분 사이에  
기계 학습 알고리즘.

물론 이 책을 읽는 방법은 두 가지 이상이다. 대부분의 독자  
하향식 및 상향식 접근 방식의 조합을 사용하여 배우며, 때로는 더 복잡한 개념을 시도하기 전에 기본  
적인 수학 기술을 구축하고 기계 응용 프로그램을 기반으로 주제를 선택합니다.

학습.

1부는 수학에 관한 것입니다.

이 책에서 다루는 머신 러닝의 4가지 기둥(그림 1.1 참조)  
파트 1에 나와 있는 견고한 수학적 기초가 필요합니다.

우리는 숫자 데이터를 벡터로 표현하고 그러한 테이블을 표현합니다.  
데이터를 행렬로. 벡터와 행렬에 대한 연구를 선형 대수학 이라고 합니다.

선형 대수학

이는 2장에서 소개합니다. 행렬로서의 벡터 모음은 다음과 같습니다.  
거기에 설명도.

실제 세계에서 두 객체를 나타내는 두 벡터가 주어졌을 때  
그들의 유사성에 대해 진술하기 위해. 아이디어는 벡터가  
유사함은 우리 기계에서 유사한 출력을 가질 것으로 예측되어야 합니다.  
학습 알고리즘(예측자). 벡터 간의 유사성 개념을 공식화하려면 두 벡터를 다음과 같이 취하는 연산을  
도입해야 합니다.

해석 기하학

유사성을 나타내는 숫자 값을 입력하고 반환합니다. 유사성과 거리의 구성은 해석 기하학 의 핵심 이  
며 다음과 같습니다.

3장에서 논의.

행렬

4장에서는 행렬과 행렬 분해에 대한 몇 가지 기본 개념을 소개합니다. 행렬에 대한 일부 연산은 매우

분해

기계 학습에 유용하며 직관적인 해석이 가능합니다.

데이터를 더 효율적으로 학습할 수 있습니다.

우리는 종종 데이터를 진정한 기본 신호의 노이즈 관찰로 간주합니다. 기계 학습을 적용하여 식별할  
수 있기를 바랍니다.

노이즈로부터의 신호. 이를 위해서는 "소음"이 의미하는 바를 정량화할 수 있는  
언어가 필요합니다. 우리는 종종 다음과 같은 예측 변수를 원합니다.

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.

## 1.2 이 책을 읽는 두 가지 방법

15

예를 들어 특정 테스트 데이터에서 예측 값에 대한 확신을 정량화하기 위해 일종의 불확실성을 표현할 수 있습니다.

가리키다. 불확실성의 정량화는 확률론 과 확률론 의 영역

6장에서 다룹니다.

기계 학습 모델을 훈련하기 위해 일반적으로 다음과 같은 매개변수를 찾습니다.

일부 성능 측정을 최대화하십시오. 많은 최적화 기술에는 기울기의 개념이 필요합니다.

솔루션을 검색합니다. 5장에서는 벡터 미적분에 대해 설명 하고 벡터 미적분 에 대해 자세히 설명합니다.

이후에 7장에서 사용하는 그래디언트 개념

함수의 최대/최소를 찾기 위한 최적화 에 대해 이야기 합니다.

최적화

## 2부는 기계 학습에 관한 것입니다.

이 책의 두 번째 부분에서는 머신 러닝의 4가지 기둥을 소개합니다.

그림 1.1과 같이. 책의 첫 부분에서 소개된 수학적 개념이 각 기둥의 기초가 되는 방법을 설명합니다.

일반적으로 챕터는 난이도(오름차순)로 정렬됩니다.

8장에서는 기계 학습의 세 가지 구성 요소를 다시 설명합니다.

(데이터, 모델 및 매개변수 추정) 수학적 방식으로. ~ 안에

또한 실험 설정을 구축하기 위한 몇 가지 지침을 제공합니다.

기계 학습 시스템에 대한 지나치게 낙관적인 평가를 방지합니다. 목표는 다음에서 잘 수행되는

예측기를 구축하는 것임을 기억하십시오

보이지 않는 데이터.

9장에서는 선형 회귀에 대해 자세히 살펴 보겠습니다 .

목적은 입력  $x \in \mathbb{R}^D$  를 해당 관찰된 함수 값  $y \in \mathbb{R}$ 에 매핑하는 함수를 찾는 것입니다. 이를 해당 함수의 레이블로 해석할 수 있습니다.

각각의 입력. 우리는 최대 우도와 최대 사후 추정을 통한 고전적 모델 피팅(모수 추정)에 대해 논의할 것입니다.

매개변수를 통합하는 베이زي안 선형 회귀뿐만 아니라

최적화하는 대신 아웃.

10장은 Fig 차원 의 두 번째 기둥 인 차원 축소 에 초점을 맞춘다.

된 저차원 표현을 찾는 것입니다. 차원 1.1 sionalality reduction의 주요 목표는 압축

절감

고차원 데이터  $x \in \mathbb{R}^D$ 의 경우보다 분석하기가 더 쉽습니다.

원본 데이터. 회귀와 달리 차원 축소는 데이터 모델링에만 관심이 있습니다.

데이터 포인트  $x$ .

11장에서는 세 번째 기둥인 밀도 추정으로 이동합니다. 밀도 추정

밀도 추정의 목적은 주어진 데이터 세트를 설명하는 확률 분포를 찾는 것입니다. 이를 위해 가우스 혼합 모델에 중점을 둘 것입니다.

목적, 그리고 우리는

이 모델. 차원 축소에서와 같이 연결된 레이블이 없습니다.

데이터 포인트  $x \in \mathbb{R}^D$ . 그러나 우리는 저차원을 추구하지 않습니다.

데이터의 표현. 대신 밀도 모델에 관심이 있습니다.

데이터를 설명하는 것입니다.

12장은 네 번째에 대한 심층적인 논의로 이 책을 마무리합니다.

분류

기동: 분류. 서포트 벡터 머신의 맥락에서 분류에 대해 논의할 것입니다. 회귀(9장)와 유사하게 입력  $x$  와 해당 레이블  $y$  가 있습니다. 그러나 레이블이 실수 값인 회귀와 달리 분류의 레이블은 정수이므로 특수

케어.

### 1.3 연습 및 피드백

우리는 대부분 펜과 종이로 할 수 있는 파트 I에서 몇 가지 연습을 제공합니다. 2부에서는 이 책에서 논의하는 기계 학습 알고리즘의 일부 속성을 탐색하기 위해 프로그래밍 자습서(jupyter 노트북)를 제공합니다.

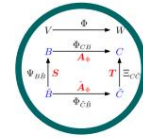
Cambridge University Press에서 이 책을 무료로 다운로드할 수 있도록 함으로써 교육과 학습을 민주화하려는 우리의 목표를 강력하게 지지해 주셔서 감사합니다.

<https://mml-book.com>

튜토리얼, 정오표 및 추가 자료를 찾을 수 있습니다. 앞의 URL을 사용하여 실수를 보고하고 피드백을 제공할 수 있습니다.

## 2

## 선형 대수학



직관적인 개념을 공식화할 때 일반적인 접근 방식은 객체(기호) 집합과 이러한 객체를 조작하는 규칙 집합을 구성하는 것입니다. 이것은 대수학 으로 알려져 있습니다. 선형 대수학은 벡터와 벡터를 조작하기 위한 특정 대수학 규칙에 대한 연구입니다. 우리 중 많은 사람들이 학교에서 알고 있는 벡터를 "기하학적 벡터"라고 하며 일반적으로 문자 위의 작은 화살표로 표시됩니다(예:  $\vec{x}$  및  $\vec{y}$ ).

이 책에서 우리는 벡터의 보다 일반적인 개념을 논의하고 굵은 글자를 사용하여 표현합니다(예:  $x$  및  $y$ ).

일반적으로 벡터는 같은 종류의 다른 객체를 생성하기 위해 함께 더하고 스칼라로 곱할 수 있는 특수 객체입니다. 추상적인 수학적 관점에서 이 두 가지 속성을 만족하는 모든 객체는 벡터로 간주될 수 있습니다. 다음은 이러한 벡터 객체의 몇 가지 예입니다.

1. 기하학적 벡터. 벡터의 이 예는 고등학교 수학 및 물리학에서 친숙할 수 있습니다. 기하 벡터 (그림 2.1(a) 참조)는 (적어도 2차원으로) 그릴 수 있는 방향 세그먼트입니다. 두 개의 기하 벡터는 또 다른 기하 벡터입니다. 또한 스칼라  $\lambda x$ 의 곱셈,  $\lambda \in \mathbb{R}$ 도 기하 벡터입니다. 사실, 이것은  $\lambda$ 로 스케일된 원래 벡터입니다. 따라서 스칼라 곱셈은 이 집합에서 닫힌 벡터 개념의 인스턴스입니다. 벡터를 기하학적 벡터로 해석하면 방향과 크기에 대한 직관을 사용하여 추상적 연산에 대해 추론할 수 있습니다.

2. 다항식도 벡터입니다. 그림 2.1(b) 참조: 두 개의 다항식은 다음을 수행할 수 있습니다.

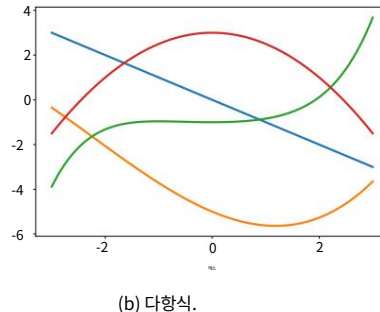
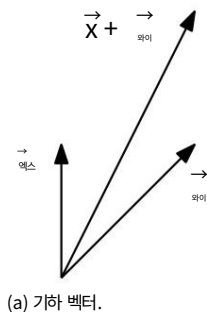


그림 2.1 벡터의 다른 유형. 벡터는 (a) 기하학적

벡터 및 (b) 다항식.

함께 더하면 다른 다항식이 생깁니다. 스칼라  $\lambda \in \mathbb{R}$  로 곱할 수 있으며 결과도 다항식입니다. 따라서 다항식은 (다소 특이한) 벡터의 인스턴스입니다.

다항식은 기하학적 벡터와 매우 다릅니다. 기하학적 벡터가 구체적인 "도면"인 반면, 다항식은 추상적 개념입니다. 그러나 이전에 설명한 의미에서 둘 다 벡터입니다.

3. 오디오 신호는 벡터입니다. 오디오 신호는 일련의 숫자로 표시됩니다. 오디오 신호를 함께 추가할 수 있으며 그 합은 새로운 오디오 신호입니다. 오디오 신호를 스케일링하면 오디오 신호도 얻습니다.

따라서 오디오 신호도 일종의 벡터입니다.

4.  $\mathbb{R}^n$  의 요소 ( $n$  실수의 튜플)는 벡터입니다.  $\mathbb{R}^n$  은 다항식보다 더 추상적이며 이 책에서 초점을 맞춘 개념입니다. 예를 들어,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3 \quad (2.1)$$

숫자의 삼중항의 예입니다. 두 벡터  $a, b \in \mathbb{R}^n$  을 구성 요소별로 더하면 다른 벡터가 생성됩니다.  $a + b = c \in \mathbb{R}^n$  . 더욱이,  $a \in \mathbb{R}^n$  에  $\lambda \in \mathbb{R}$  을 곱하면 크기가 조정된 벡터  $\lambda a \in \mathbb{R}^n$  이 됩니다.

확인해 주십시오  
컴퓨터에서 구현할  
때 배열 연산이 실제로 벡  
터 연산을 수행하는지 여  
부.

벡터를  $\mathbb{R}^n$  의 요소로 고려 하면 컴퓨터의 실수 배열에 느슨하게 대응한다는 추가적인 이점이 있습니다. 많은 프로그래밍 언어는 벡터 연산을 포함하는 알고리즘의 편리한 구현을 허용하는 배열 연산을 지원합니다.

선형에 대한 Pavel  
Grinfeld의 시리즈  
대수학:  
<http://tinyurl.com/nahclwm>  
Gilbert Strang의  
선형 과정  
대수학:  
<http://tinyurl.com/29p5q8j>  
3Blue1Brown 시리즈  
선형 대수: <https://tinyurl.com/h5g4kps>

선형 대수학은 이러한 벡터 개념 간의 유사성에 초점을 맞춥니다.

그것들을 더하고 스칼라로 곱할 수 있습니다. 선형 대수학의 대부분의 알고리즘은  $\mathbb{R}^n$ 에서 mulated되기 때문에 우리는 주로  $\mathbb{R}^n$  의 벡터에 초점을 맞출 것입니다. 8장에서 데이터가  $\mathbb{R}^n$  에서 벡터로 표현되는 경우가 많다는 것을 알게 될 것입니다.

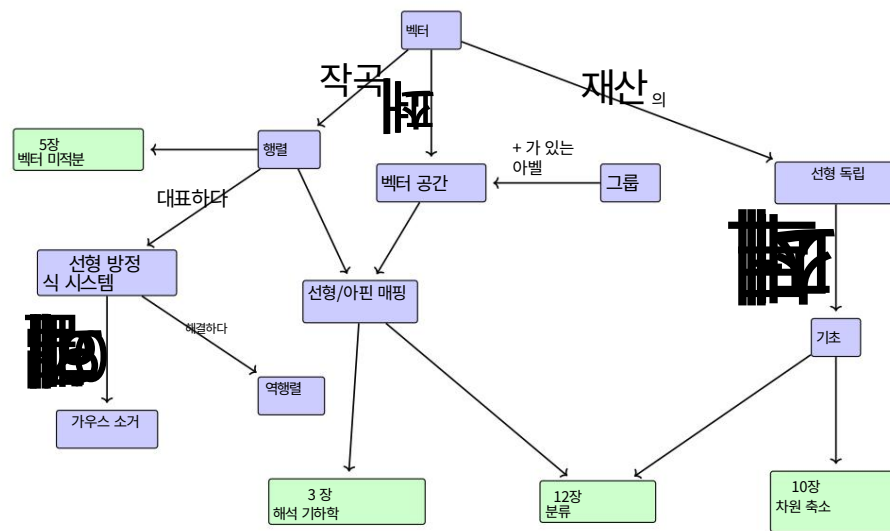
이 책에서는 모든 종류의 벡터와  $\mathbb{R}^n$  사이에 1:1 대응이 있는

유한 차원 벡터 공간에 초점을 맞출 것입니다. 편리한 경우 기하학적 벡터에 대한 직관을 사용하고 배열 기반 알고리즘을 고려합니다.

수학의 주요 개념 중 하나는 "닫힘" 개념입니다. 이것은 질문입니다. 내가 제안한 작업으로 인해 발생할 수 있는 모든 것의 집합은 무엇입니까? 벡터의 경우: 작은 벡터 집합으로 시작하여 서로 더하고 크기를 조정하면 얻을 수 있는 벡터 집합은 무엇입니까? 그 결과 벡터 공간이 생성됩니다(2.4절). 벡터 공간의 개념과 그 속성은 많은 기계 학습의 기초가 됩니다. 이 장에서 소개된 개념은 그림 2.2에 요약되어 있습니다.

이 장은 주로 Drumm and Weil(2001), Strang(2003), Hogben(2013), Liesen and Mehrmann(2015), Pavel Grinfeld의 Linear Algebra 시리즈의 강의 노트와 책을 기반으로 합니다. 기타 우수

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.



총

$$a_{i1}x_1 + \cdots + a_{in}x_n \quad (2.2)$$

자원  $R_i$  의 많은 단위 . 최적의 생산 계획  $(x_1, \dots, x_n) \in \mathbb{R}^n$  ,  
따라서 다음 방정식 시스템을 충족해야 합니다.

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \quad (2.3)$$

여기서  $a_{ij} \in \mathbb{R}$  및  $b_i \in \mathbb{R}$ .

선형 방정식 시스템

식 (2.3)은 선형 방정식 시스템의 일반적인 형태이며 ,  
 $x_1, \dots, x_n$  은 이 시스템 의 미지수 입니다. 모든  $n$ -튜플  $(x_1, \dots, x_n) \in \mathbb{R}^n$  (2.3)을 만족하는 것은 선형 방정식 시스템의 솔루션 입니다.

해결책

## 예 2.2

선형 방정식 시스템

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 & x_1 - x_2 + 2x_3 &= 2 & (1) \\ 2x_1 + 3x_3 &= 1 & & & (2) \end{aligned} \quad (2.4)$$

(삼)

해가 없습니다. 처음 두 방정식을 더하면  $2x_1 + 3x_3 = 5$ 가 됩니다.  
세 번째 방정식 (3)과 모순됩니다.

선형 방정식 시스템을 살펴보겠습니다.

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 & x_1 - x_2 + 2x_3 &= 2 & (1) \\ x_2 + x_3 &= 2 & & & (2) \end{aligned} \quad (2.5)$$

(삼)

첫 번째 및 세 번째 방정식에서  $x_1 = 1$  이 됩니다. (1)+(2)에서,  
 $2x_1 + 3x_3 = 5$ , 즉  $x_3 = 1$ 을 얻습니다. (3)에서  $x_2 = 1$  을 얻습니다 .  
따라서  $(1, 1, 1)$  은 유일하게 가능하고 고유한 솔루션입니다 (다음을 확인 하십시오.  
 $(1, 1, 1)$  은 플러그인에 의한 솔루션입니다).

세 번째 예로 다음을 고려합니다.

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 & x_1 - x_2 + 2x_3 &= 2 & (1) \\ 2x_1 + 3x_3 &= 5 & & & (2) \end{aligned} \quad (2.6)$$

(삼)

(1)+(2)=(3)이므로 세 번째 방정식(중복)을 생략할 수 있습니다. 에서  
(1)과 (2)에서 우리는  $2x_1 = 5 - 3x_3$  및  $2x_2 = 1 + x_3$ 을 얻습니다. 우리는  $x_3 = a \in \mathbb{R}$  을 정의합니다.  
임의의 트리플렛과 같은 자유 변수로

$$\frac{5}{2} - \frac{3a}{2}, \frac{1}{2} + \frac{a}{2} \text{에이}, \quad \in \mathbb{R} \quad (2.7)$$



## 2.1 선형 방정식 시스템

21

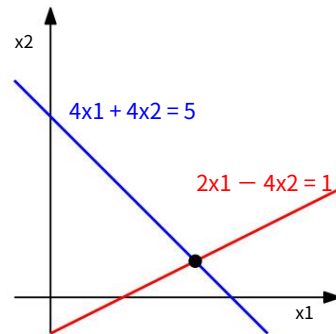


그림 2.3 두 개의 변수가 있는 두 개의 선형 방정식 시스템의 해 공간은 두 변수의 교집합으로 기하학적으로 해석될 수 있습니다.

윤곽. 모든 선형 방정식은 다음을 나타냅니다. 라인.

선형 방정식 시스템의 솔루션입니다. 즉, 무한히 많은 솔루션을 포함하는 솔루션 세트를 얻습니다.

일반적으로 선형 방정식의 실수 값 시스템에 대해 아니요, 정확히 하나 또는 무한히 많은 솔루션을 얻습니다. 선형 회귀(9장)는 선형 방정식 시스템을 풀 수 없을 때 예제 2.1의 버전을 풉니다.

비고 (선형 방정식 시스템의 기하학적 해석). 두 개의 변수  $x_1, x_2$ 가 있는 선형 방정식 시스템에서 각 선형 방정식은  $x_1x_2$  평면의 선을 정의합니다. 선형 연립방정식의 해는 모든 방정식을 동시에 만족해야 하므로 해 세트는 이 선의 교집합입니다. 이 교차 세트는 선(선형 방정식이 동일한 선을 설명하는 경우), 점 또는 비어 있는(선이 평행한 경우)일 수 있습니다. 시스템에 대한 그림이 그림 2.3에 나와 있습니다.

$$\begin{aligned} 4x_1 + 4x_2 &= 5 \\ 2x_1 - 4x_2 &= 1 \end{aligned} \quad (2.8)$$

여기서 솔루션 공간은 점  $(x_1, x_2) = (1, )$ 입니다. 마찬가지로<sup>1</sup>3개의 변수에 대해 각 선형 방정식은 3차원 공간의 평면을 결정합니다. 이러한 평면을 교차할 때, 즉 모든 선형 방정식을 동시에 만족할 때 평면, 선, 점 또는 비어 있는 솔루션 세트를 얻을 수 있습니다(평면에 공통 교차점이 없을 때). ◆

선형 방정식 시스템을 풀기 위한 체계적인 접근을 위해 유용한 압축 표기법을 소개합니다. 계수  $a_{ij}$ 를 벡터로 수집하고 벡터를 행렬로 수집합니다. 즉, (2.3)의 시스템을 다음 형식으로 작성합니다.

$$\begin{array}{ccccccc} a_{11} & & a_{12} & & \text{에이엔} & & b_1 \\ \vdots & x_1 + & \vdots & x_2 + \cdots + & \vdots & x_n = & \vdots \\ \text{오전1} & & \text{오전2} & & \text{암} & & \text{비엠} \end{array} \quad (2.9)$$

$$\Leftrightarrow \begin{pmatrix} a_{11} & \cdots & 1n \\ a_{m1} & \cdots & amn \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

## 2.2 행렬

행렬

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & 1n \\ a_{21} & a_{22} & \cdots & 2n \end{pmatrix}, a_{ij} \in \mathbb{R}.$$

오전 1시 2시  $\cdots$  오전

열

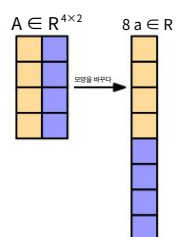
열

행 벡터

열 벡터

그림 2.4 열을  
쌓아서 행렬 A  
를 긴 벡터 a로 나타낼  
수 있습니다.

### 2.2.1 행렬 덧셈과 곱셈



사이즈 참고하세요

행렬.

C =

`np.einsum('il, lj',  
A, B)`

$$A + B := \begin{pmatrix} a_{11} + b_{11} & \cdots & 1n + b_{1n} \\ a_{m1} + b_{m1} & \cdots & amn + b_{mn} \end{pmatrix} \quad m \times n.$$

$$B \in \mathbb{R}^{n \times k}$$

$$c_{ij} = \sum_{l=1}^n a_{il} b_{lj}, \quad \text{이. 나} = 1, \dots, m, j = 1, \dots, \text{케}$$

## 2.2 행렬

23

A 및 n 행에서  
B 우리가 할 수 있도록  
l = 1, 에 대해  $a_{il}b_{lj}$   
를 계산합니다. ...,  
일반적으로 두 벡터  
a, b 사이의 내적은  
>b 또는  $h_{a, b}$ 로 표  
시됩니다.

$$\begin{matrix} \vdots \\ \vdots \\ \vdots \end{matrix} = \begin{matrix} \vdots \\ \vdots \\ \vdots \end{matrix} \\ n \times m$$

## ◆ 하다마르 제품

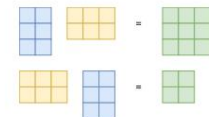
$$\begin{array}{l} \begin{matrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{matrix} \in \mathbb{R}^{2 \times 3} \quad B = \begin{matrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{matrix} \in \mathbb{R}^{3 \times 2} \\ \\ AB = \begin{matrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{matrix} \begin{matrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{matrix} = \begin{matrix} 2 & 3 \\ 2 & 5 \end{matrix} \quad 2 \times 2, \\ \\ \text{학사} = \begin{matrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{matrix} \begin{matrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{matrix} = \begin{matrix} 6 & 4 & 2 \\ -2 & 0 & 2 \\ 3 & 2 & 1 \end{matrix} \quad 3 \times 3. \end{array}$$

그림 2.5 두 행렬 모두

곱셈 AB와 BA는

정의된 차수

결과가 다를 수 있  
습니다.



단위 행렬

$$\begin{matrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{matrix} \quad n \times n$$

에서 :=

대각선에 1 을 포함 하고 다른 모든 곳에 0 을 포함 하는  $n \times n$ -행렬로 .

이제 행렬 곱셈, 행렬 덧셈 및  
단위 행렬, 행렬의 몇 가지 속성을 살펴보겠습니다.

연관성

■ 연관성:

$$\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times q} : (AB)C = A(BC) \quad (2.18)$$

분배성

■ 분포:

$$\forall A, B \in \mathbb{R}^{m \times n}, C, D \in \mathbb{R}^{n \times p} : (A + B)C = AC + BC \quad (2.19a)$$

$$A(C + D) = AC + AD \quad (2.19b)$$

■ 단위 행렬을 사용한 곱셈:

$$\forall A \in \mathbb{R}^{m \times n} : I_m A = A I_n = A \quad (2.20)$$

$m = n$  에 대해  $I_m = I_n$  에 유의하십시오 .

### 2.2.2 역과 전치

정방 행렬은 동일  
한 수의 열과 행을 갖  
습니다.

정의 2.3 (역). 정방 행렬  $A \in \mathbb{R}^{n \times n}$  을 고려하십시오 . 행렬  $B \in \mathbb{R}^{n \times n}$  이  $AB = I_n = BA$  라는  
속성을 갖습니다 .  $B$  는  $A$ 의 역함수 라고 하며  $A$  로 표시됩니다 .

역

불행히도, 모든 행렬  $A$  가 역행렬을 가지는 것은 아닙니다 . 역행렬이 존재하지 않습<sup>-1</sup> . 이 경우  
니다.  $A$  는 정규/역/비특이, 그렇지 않으면 특이/비가역이라고 합니다. 역행렬이 존재할 때 고유합니  
다. 2.3절에서 우리는 선형 방정식 시스템을 풀어서 행렬의 역행렬을 계산하는 일반적인 방법을 논의  
할 것입니다.

정기적인  
뒤집을 수 있는  
비단수 단수

비가역

비고 ( $2 \times 2$  행렬의 역행렬의 존재). 행렬을 고려하십시오

$$A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2 \times 2}. \quad (2.21)$$

$A$  를 곱하면

$$I^0 := \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \quad (2.22)$$

우리는 얻는다

$$AA^0 = \begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} = (a_{11}a_{22} - a_{12}a_{21})I. \quad (2.23)$$

그러므로,

$$I^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \quad (2.24)$$

$a_{11}a_{22} - a_{12}a_{21} \neq 0$ 인 경우에만 . 섹션 4.1에서  $a_{11}a_{22} -$

$a_{12}a_{21}$  은  $2 \times 2$  행렬의 행렬식입니다. 게다가, 우리는 일반적으로 행렬이 역행성인지 여부를 확인하기 위해 행렬식을 사용할 수 있습니다.

예제 2.4(역행렬)  
행렬

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 4 & 5 \\ 6 & 7 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} -7 & -7 & 6 \\ 2 & 1 & -1 \\ 4 & 5 & -4 \end{pmatrix} \quad (2.25)$$

$AB = I = BA$  이므로 서로 역입니다.

정의 2.4 (전치).  $A \in \mathbb{R}^{m \times n}$  에 대해 행렬  $B \in \mathbb{R}^{n \times m}$

$b_{ij} = a_{ji}$  는  $A$ 의 전치 라고 합니다. 우리는  $B = A^T$  라고 씁니다.

일반적으로  $A^T A$  의 열을 행 으로 쓰면 얻을 수 있습니다.  $A^T A$  의 열은 행 으로 쓰면 얻을 수 있습니다.  $A^T A$  의 열은 행 으로 쓰면 얻을 수 있습니다.

바꾸어 놓다

행렬  $A$ 의 주대각선(때로는 "주대각선", "주대각선", "선행대각선" 또는 "주대각선"이라고도 함)은 항목 모음입니다.

$$AA^T = I = A^T A \quad (2.26)$$

$$(AB)^T = B^T A^T \quad (2.27)$$

$$(A + B)^T = A^T + B^T \quad (2.28)$$

$$(A^T)^T = A \quad (2.29)$$

$$(A + B)^T = A^T + B^T \quad (2.30)$$

$$(AB)^T = B^T A^T \quad (2.31)$$

$A_{ij}$  여기서  $i = j$ .

(2.28)의 스칼라 경우는 다음

과 같습니다.

$$\overline{12+4} = \overline{1} \overline{6} = \overline{1} \overline{2} + \overline{1} \overline{4}$$

정의 2.5 (대칭 행렬).  $A$  행렬  $A \in \mathbb{R}^{n \times n}$  은 대칭 행렬  $A = A^T$  인 경우 대칭입니다.

$(n, n)$ -행렬 만 대칭일 수 있습니다. 일반적으로  $(n, n)$ -행렬 은 행과 열의 동일한 수의 제곱 행렬을 가지므로 정방 행렬 이라고도 합니다. 또한  $A$  가 가역이면  $A^{-1}$  도 마찬가지로 대칭입니다 ( $A$  비고 (대칭 행렬 의 합과 곱 ). 대칭 행렬  $A, B \in \mathbb{R}^{n \times n}$  의 합 은 항상 대칭입니다. 그러나 곱 은 항상 정의되지만 일반적으로 대칭이 아닙니다.

$$(A+B)^T = (A^T + B^T)^T = A + B$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad (2.32)$$



### 2.2.3 스칼라에 의한 곱셈

행렬에 스칼라  $\lambda \in \mathbb{R}$  을 곱할 때 행렬에 어떤 일이 발생하는지 살펴 보겠습니다.  $A \in \mathbb{R}^{m \times n}$  및  $\lambda \in \mathbb{R}$  이라고 합시다. 그러면  $\lambda A = K$ ,  $K_{ij} = \lambda a_{ij}$  입니다.

실제로  $\lambda$  는  $A$  의 각 요소를 스케일링합니다.  $\lambda, \psi \in \mathbb{R}$  에 대해 다음이 성립합니다.

연관성

■

■

,

분배성

$$\lambda(C) = C = \lambda C$$

$$\text{분포: } (\lambda + \psi)C = \lambda C + \psi C, C \in \mathbb{R}^{m \times n} \quad \lambda(B + C) = \lambda B + \lambda C, B, C \in \mathbb{R}^{m \times n}$$

$$\text{예} := \begin{pmatrix} 1 & 2 & 3 \\ 4 & & \end{pmatrix},$$

$$\begin{aligned} (\lambda + \psi)C &= \begin{pmatrix} \lambda + \psi & 2\lambda + 2\psi & 3\lambda + 3\psi \\ 4\lambda + 4\psi & & \end{pmatrix} \\ &= \begin{pmatrix} \lambda & 2\lambda & 3\lambda \\ 4\lambda & & \end{pmatrix} + \begin{pmatrix} \psi & 2\psi & 3\psi \\ 4\psi & & \end{pmatrix} = \lambda C + \psi C. \end{aligned}$$

## 2.2.4 선형 방정식 시스템의 간결한 표현

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 1 \\ 4x_1 - 2x_2 - 7x_3 &= 8 \\ 9x_1 + 5x_2 - 3x_3 &= 2 \end{aligned}$$

$$\begin{pmatrix} 2 & 3 & 5 \\ 4 & -2 & -7 \\ 9 & 5 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 2 \end{pmatrix}$$

## 2.3 선형 방정식의 풀이 시스템

27

## 2.3 선형 방정식의 풀이 시스템

(2.3)에서 우리는 방정식 시스템의 일반적인 형태를 소개했습니다. 즉,

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \quad (2.37)$$

여기서  $a_{ij} \in \mathbb{R}$  및  $b_i \in \mathbb{R}$ 은 알려진 상수이고  $x_j$ 는 미지수,  $n$ 입니다. 지금까지 행렬을  $Ax = b$ 를 쓰는 행렬의 선형 방정식 시스템을 풀 수 있는 방법을 소개했습니다. (2.37)에서, 우리는 선형 방정식의 시스템을 푸는 데 초점을 맞추고 행렬의 역행렬을 찾는 알고리즘을 제공합니다.

## 2.3.1 특정 및 일반 솔루션

일반적으로 선형 연립방정식을 푸는 방법을 논의하기 전에 예를 살펴보겠습니다. 연립방정식을 고려하라

$$\begin{pmatrix} 1 & 0 & 8 & -4 & 0 & 1 & 2 & 12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 42 \\ 8 \end{pmatrix}. \quad (2.38)$$

시스템에는 2개의 방정식과 4개의 미지수가 있습니다. 따라서 일반적으로 우리는 무한히 많은 솔루션을 기대합니다. 이 방정식 시스템은 처음 두 열이 1과 0으로 구성된 특히 쉬운 형식입니다. 스칼라  $x_1, \dots, x_4$ ,  $Pb$ 가 (2.38)의 우변이 되도록 합니다. (2.38)의 문제에 대한 해는 첫 번째 열의 42배와 두 번째 열의 8배를 취하여 즉시 찾을 수 있습니다.  $i=1$   $x_i c_i = b$ , 여기서  $c_i$ 를 행렬의  $i$ 번째 열로 정의하고

$$b = \begin{pmatrix} 42 \\ 8 \end{pmatrix} = 42 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 8 \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.39)$$

따라서 해는  $[42, 8, 0, 0]^T$ 입니다. 이 솔루션을 특정 특정 솔루션 솔루션 또는 특수 솔루션이라고 합니다. 그러나 이것이 선형 방정식 의 이 특수 솔루션 시스템의 유일한 솔루션은 아닙니다. 다른 모든 솔루션을 캡처하려면 행렬의 열을 사용하여 간단하지 않은 방식으로 0을 생성하는 데 창의적이어야 합니다. 특수 솔루션에 0을 추가해도 특수 솔루션은 변경되지 않습니다. 그렇게 하기 위해 처음 두 개의 열을 사용하여 세 번째 열을 표현합니다(이는 매우 간단한 형식임).

$$\begin{pmatrix} 8 \\ 2 \end{pmatrix} = 8 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.40)$$

그래서  $0 = 8c_1 + 2c_2 - 1c_3 + 0c_4$  및  $(x_1, x_2, x_3, x_4) = (8, 2, -1, 0)$ 입니다. 사실, 이 해를  $\lambda_1 \in \mathbb{R}$ 만큼 스케일링하면 0 벡터가 생성됩니다. 즉,

$$\begin{array}{ccccccc} & & & & 8 & & \\ & & & & 2 & & \\ 1 & 0 & 8 & -4 & & & \\ 0 & 1 & 2 & 12 & & & \\ & & & & -1 & & \\ & & & & 0 & & \end{array} \quad \lambda_1 \quad = \lambda_1(8c_1 + 2c_2 - c_3) = 0. \quad (2.41)$$

동일한 추론 라인에 따라 처음 두 열을 사용하여 (2.38)에서 행렬의 네 번째 열을 표현하고 다음과 같이 0의 중요하지 않은 버전 집합을 생성 합니다.

$$\begin{array}{ccccccc} & & & & -4 & & \\ & & & & 12 & & \\ 1 & 0 & 8 & -4 & 0 & 1 & \\ 2 & 12 & & & 0 & & \\ & & & & -1 & & \end{array} \quad \lambda_2 \quad = \lambda_2(-4c_1 + 12c_2 - c_4) = 0 \quad (2.42)$$

일반 솔루션

모든  $\lambda_2 \in \mathbb{R}$ 에 대해. 모든 것을 종합하면 일반 솔루션 이라고 하는 (2.38)의 방정식 시스템의 모든 솔루션을 세트 로 얻습니다.

$$x \in \mathbb{R}^4 : x = \begin{array}{ccc} 42 & 8 & -4 \\ 8 & 2 & 12 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{array} + \lambda_1 \begin{array}{ccc} 8 & 2 & -1 \\ 2 & 12 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} + \lambda_2 \begin{array}{ccc} -4 & 12 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}, \lambda_1, \lambda_2 \in \mathbb{R} \quad (2.43)$$

주목. 우리가 따랐던 일반적인 접근 방식은 다음 세 단계로 구성되었습니다.

1.  $Ax = b$ 에 대한 특정 솔루션을 찾습니다.
2.  $Ax = 0$ 에 대한 모든 솔루션을 찾습니다.
3. 1.과 2.의 솔루션을 일반 솔루션에 결합합니다.

일반 솔루션도 특정 솔루션도 고유하지 않습니다. ♦

앞의 예에서 선형 방정식의 시스템은 (2.38)의 행렬이 특히 편리한 형식을 가지고 있기 때문에 풀기 쉬웠고, 이를 통해 검사를 통해 특정 솔루션과 일반 솔루션을 찾을 수 있었습니다. 그러나 일반 방정식 시스템은 이러한 단순한 형태가 아닙니다.

다행스럽게도 선형 방정식 시스템을 가우스 소거법이라는 간단한 형식으로 변환하는 건설적인 알고리즘 방식이 있습니다. 가우스안 제거의 핵심은 방정식 시스템을 간단한 형태로 변환하는 선형 방정식 시스템의 기본 변환입니다. 그런 다음 (2.38)의 예제 컨텍스트에서 방금 논의한 간단한 형식에 세 단계를 적용할 수 있습니다.

### 2.3.2 기본 변환 선형 방정식 시스템을

기본 변환

푸는 열쇠는 솔루션 세트를 동일하게 유지하지만 방정식 시스템을 더 간단한 형태로 변환하는 기본 변환입니다.



## 2.3 선형 방정식의 풀이 시스템

29

- 두 방정식의 교환(시스템을 나타내는 행렬의 행 방정식)
- 상수  $\lambda \in \mathbb{R} \setminus \{0\}$  이 있는 방정식(행)의 곱셈
- 두 방정식(행)의 덧셈

## 예 2.6

$\in \mathbb{R}$ 에 대해 다음 연립방정식의 모든 해를 구합니다.

$$\begin{aligned} -2x_1 + 4x_2 - 2x_3 - x_4 + 4x_5 &= -3 \\ 4x_1 - 8x_2 + 3x_3 - 3x_4 + x_5 &= 2 \\ x_1 - 2x_2 + x_3 - x_4 + x_5 &= 0 \\ x_1 - 2x_2 - 3x_4 + 4x_5 &= a \end{aligned} \quad (2.44)$$

우리는 이 방정식 시스템을 콤팩트 행렬로 변환하는 것으로 시작합니다. 표기법  $Ax = b$ . 우리는 더 이상 변수  $x$  를 명시적으로 언급하지 않으며 증대 행렬 을 작성 합니다( $A | b$  형식).

증강 행렬

$$\begin{array}{cccc|c} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{array} \quad \begin{array}{l} R_3 \text{ 로 교체} \\ R_1 \text{ 으로 교체} \end{array}$$

여기서 우리는 수직선을 사용하여 왼쪽과 오른쪽을 구분했습니다.

(2.44)의 오른쪽. 의 변환을 나타내기 위해 사용합니다.

기본 변환을 사용하는 증강 행렬.

1 행 과 3 행 을 바꾸면

$$\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 4 & -8 & 3 & -3 & 4 & 2 \\ 1 & -2 & 0 & -3 & 4 & a \end{array} \quad \begin{array}{l} -4R_1 \\ +2R_1 \\ -R_1 \end{array}$$

증강  
매트릭스  $A | b$   
컴팩트하게  
를 나타냅니다  
선형 시스템  
방정식  $Ax = b$ .

이제 표시된 변환을 적용할 때(예: 행 1 빼기  
행 2)에서 네 번, 우리는

$$\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{array} \quad \begin{array}{l} -R_2 - R_3 \\ \cdot (-1) \\ \cdot (-\frac{1}{3}) \\ 0 + 1 \end{array}$$

$$\begin{array}{cccc|c} 1 & -2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \quad \begin{array}{l} 3-2 \\ 0+1 \end{array}$$

행-제대 형식

이 (증강) 행렬은 편리한 형태인 행-사다리꼴 형태 (REF)입니다. 이 압축 표기법을 우리가 찾는 변수를 사용하여 명시적 표기법으로 되돌리면 다음을 얻습니다.

$$\begin{aligned} x_1 - 2x_2 + x_3 - x_4 + x_5 &= 0 \\ x_3 - x_4 + 3x_5 &= -2 \\ x_4 - 2x_5 &= 1 \end{aligned} \quad \text{. (2.45)}$$

특정 솔루션

$= -1$ 인 경우에만 이 시스템을 풀 수 있습니다. 특정 솔루션은

$$\begin{aligned} x_1 &= 2 \\ x_2 &= 0 \\ x_3 &= -1 \\ x_4 &= 1 \\ x_5 &= 0 \end{aligned} \quad \text{. (2.46)}$$

일반 솔루션

가능한 모든 솔루션 세트를 캡처하는 일반 솔루션은 다음과 같습니다.

$$x \in \mathbb{R}^5 : x = \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \quad \text{. (2.47)}$$

다음에서 우리는 선형 방정식 시스템의 특정하고 일반적인 솔루션을 얻는 건설적인 방법을 자세히 설명합니다.

피벗

비고 (피벗 및 계단 구조). 행의 선행 계수(왼쪽에서 0이 아닌 첫 번째 숫자)를 피벗 이라고 하며 항상 그 위에 있는 행의 피벗 오른쪽에 있습니다. 따라서 행-사각형 형식의 모든 방정식 시스템은 항상 "계단" 구조를 갖습니다. ♦

행-제대 형식

정의 2.6 (행 계층 형식). 행렬은 다음과 같은 경우 행-사각형 형식입니다.

- 0만 포함하는 모든 행은 행렬의 맨 아래에 있습니다. 따라서 0이 아닌 요소를 하나 이상 포함하는 모든 행은 0만 포함하는 행의 맨 위에 있습니다.
- 0이 아닌 행만 보면 왼쪽에서 0이 아닌 첫 번째 숫자( 피벗 또는 선행 계수라고도 함)는 항상 그 위에 있는 행의 피벗 오른쪽에 있습니다.

피벗

선행 계수 다른 텍스트에서는 피벗이 1이어야 하는 경우가 있습니다.

기본 변수

자유 변수

비고 (기본 및 자유변수). 행-사다리꼴 형태의 피벗에 해당하는 변수를 기본 변수 라고 하고 나머지 변수를 자유 변수라고 합니다. 예를 들어, (2.45)에서  $x_1, x_3, x_4$  는 기본 변수인 반면  $x_2, x_5$  는 자유 변수입니다.

비고 (특정 솔루션 얻기). 행-사각형 형태는

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.

특정 솔루션을 결정해야 할 때 삶이 더 쉬워집니다. 이를 위해  $b = PP^i = 1$ , 가 되도록 피벗 열을 사용하여 방정식 시스템의 우변을 표현합니다.  $\dots$ ,  $P$ 는 피벗 열입니다.  $\lambda_1, \lambda_2, \lambda_3$ 는, 가장 오른쪽 피벗 열에서 결정됩니다.

이전 예에서 우리는  $\lambda_1, \lambda_2, \lambda_3$ 를 찾으려고 노력할 것입니다.

$$\begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.48)$$

여기에서  $\lambda_3 = 1, \lambda_2 = -1, \lambda_1 = 2$ 를 비교적 직접적으로 알 수 있습니다. 모든 것을 통합할 때 계수를 암시적으로 0으로 설정한 비피벗 열을 잊어서는 안 됩니다. 따라서 다음을 얻습니다. 특정 솔루션  $x = [2, 0, -1, 1, 0]^T$ . ♦

비고 (Reduced Row Echelon Form). 방정식 시스템은 행-사각형 형식인 경우 기약 된 행-사각형 형식 (동어: 행 기약형 또는 행 정준 형식)입니다.

- 행 사선 형식입니다.
- 모든 피벗은 1입니다.
- 피벗은 해당 열에서 0이 아닌 유일한 항목입니다.

♦

축소된 행-사각형 형식은 선형 방정식 시스템의 일반 솔루션을 간단한 방법으로 결정할 수 있게 해주기 때문에 섹션 2.3.3의 뒷부분에서 중요한 역할을 할 것입니다.

가우스

비고 (가우스 제거). 가우스 소거법 은 소거법 이 기본 변환을 수행하여 선형 방정식 시스템을 축소된 행-사각형 형태로 만드는 알고리즘입니다. ♦

실시예 2.7(축소행 사다리꼴 형태)

다음 행렬이 축소된 행-사각형 형식인지 확인합니다(피벗은 굵게 표시됨).

$$A = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad (2.49)$$

$Ax = 0$ 의 솔루션을 찾는 핵심 아이디어는 피벗 열의 (선형) 조합으로 표현해야 하는 비피벗 열을 보는 것입니다. 축소된 행 사다리꼴 형식은 이것을 비교적 간단하게 만들고 왼쪽에 있는 피벗 열의 합과 배수로 비피벗 열을 표현합니다. 두 번째 열은 첫 번째 열의 3 배입니다 (피벗은 무시할 수 있습니다. 두 번째 열의 오른쪽에 있는 열). 따라서 0을 얻으려면 빼야 합니다.



핵심  
널스페이스

$$A = \begin{pmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{pmatrix}$$
$$\sim = \begin{pmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$
$$: x = \lambda_1 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 9 \\ -4 \\ -1 \\ 0 \end{pmatrix}, \lambda_1, \lambda_2 \in \mathbb{R}$$

역수 계산  
-1

$A$ 인 ... 인 $A^{-1}$  .

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\begin{array}{cccc|ccccccc} 1 & 0 & 2 & 0 & 1 & 1 & 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & & & & & & 0 & 0 & 0 & 0 & 1 & & & \end{array}$$

$$\begin{array}{cccc|cccc} 1 & 0 & 0 & -12 & -22 & & & \\ 0 & 1 & 0 & 0 & 1 & -12 & -2 & \\ 0 & 0 & 1 & 0 & 1 & -11 & -1 & \\ 0 & 0 & 0 & & & 1 & -10 & -12 \end{array},$$

$$I^{-1} = \begin{pmatrix} -12 & -22 \\ 1 & -12 & -2 \\ 1 & -11 & -1 \\ -10 & -12 \end{pmatrix}.$$

### 2.3.4 선형 연립방정식을 풀기 위한 알고리즘

-1

-1

$$Ax = b \iff A^{-1}Ax = A^{-1}b \iff x = (A^{-1}A)^{-1}A^{-1}b$$

"기계 학습을 위한 수학" 초안(2022-01-11), 피드백: <https://mml-book.com>.

$Ax = b$  를 푸는 Moore-Penrose 의사 역행렬 ( $A > A$ ) -1A 솔루션(2.59) Moore-Penrose 의사 역행렬 을 결정하  
 을 사용합니다 . 이는 최소 노름 최소제곱해에도 해당합니다. 이 접근 방식의 단점은 행렬-행렬 곱과 기 위해  
 $A > A$  의 역합수를 계산하는 데 많은 계산이 필요하다는 것입니다 . 더욱이, 수치적 정밀도의 이유로  
 일반적으로 역 또는 의사 역을 계산하는 것은 권장되지 않습니다.

따라서 다음에서는 선형 방정식 시스템을 풀기 위한 대안적인 접근 방식에 대해 간략하게 논의합니다.

가우스 소거법은 행렬식을 계산할 때(4.1절), 벡터 집합이 선형 독립인지 여부를 확인하고(2.5  
 절), 역행렬을 계산하고(2.2.2절), 행렬의 순위를 계산할 때(절) 중요한 역할을 합니다. 섹션 2.6.2)  
 및 벡터 공간의 기초 결정(섹션 2.6.1). 가우스 소거법은 수천 개의 변수가 있는 선형 방정식 시스템  
 을 푸는 직관적이고 건설적인 방법입니다. 그러나 수백만 개의 변수가 있는 시스템의 경우 필요한 산  
 술 연산 수가 연립 방정식의 수에서 3차로 확장되기 때문에 비실용적입니다.

실제로, 많은 선형 방정식의 시스템은 Richardson 방법, Ja cobi 방법, Gauß-Seidel 방법  
 및 연속적 과이완 방법과 같은 고정 반복 방법 또는 Krylov 부분 공간 방법에 의해 간접적으로 해결  
 됩니다. 켈레 기율기, 일반화 최소 잔차 또는 이중 켈레 기율기로. 자세한 내용은 Stoer and  
 Burlirsch(2002), Strang(2003), Liesen and Mehrmann(2015)의 책을 참조하십시오.

$x^*$  를  $Ax = b$  의 해라고 하자 . 이러한 반복적인 방법의 핵심 아이디어  
 형식의 반복을 설정하는 것입니다

$$x^{(k+1)} = Cx^{(k)} + d \quad (2.60)$$

모든 반복에서 잔류 오차  $\|x^{(k+1)} - x^*\|$  를 줄이고  $x^*$ 로 수렴하는 적절한  $C$  및  $d$  에 대해 . 섹션 3.1  
 에서 벡터 간의 유사성을 계산할 수 있는 규범  $k \cdot k$ 를 소개합니다.

## 2.4 벡터 공간 지금까지 선

형 방정식 시스템과 이를 푸는 방법을 살펴보았습니다(2.3절). 우리는 선형 방정식 시스템이 행렬-벡  
 터 표기법(2.10)을 사용하여 간결하게 표현될 수 있음을 보았습니다. 다음에서는 벡터 공간, 즉 벡터  
 가 존재하는 구조화된 공간에 대해 자세히 살펴보겠습니다.

이 장의 시작 부분에서 벡터를 함께 더하고 스칼라로 곱할 수 있는 객체로 비공식적으로 특성화했  
 으며 동일한 유형의 객체로 남아 있습니다. 이제 이것을 공식화할 준비가 되었으며, 집합의 일부 구조  
 를 그대로 유지하는 이러한 요소에 대해 정의된 작업 및 요소 집합인 그룹의 개념을 소개하는 것으로  
 시작하겠습니다.

## 2.4.1 그룹

그룹은 컴퓨터 과학에서 중요한 역할을 합니다. 세트에 대한 작업을 위한 기본 프레임워크를 제공하는 것 외에도 암호화, 코딩 이론 및 그래픽에서 많이 사용됩니다.

그룹  
폐쇄  
연관성  
중성 요소  
역 요소

정의 2.7 (그룹). 집합  $G$  와 연산  $\otimes : G \times G \rightarrow G$  가  $G$  에 정의되어 있다고 가정합니다. 그러면  $G := (G, \otimes)$  는 다음이 성립하는 경우 그룹 이라고 합니다.

1.  $\otimes$  아래 의  $G$  폐쇄 :  $\forall x, y \in G : x \otimes y \in G$
2. 결합성 :  $\forall x, y, z \in G : (x \otimes y) \otimes z = x \otimes (y \otimes z)$
3. 중성 요소:  $\exists e \in G \forall x \in G : x \otimes e = x$  및  $e \otimes x = x$
4. 역 요소:  $\forall x \in G \exists y \in G : x \otimes y = e$  및  $y \otimes x = e$ , 여기서  $e$  는 역 요소를 나타냅니다.  
중립적 요소. 우리는 종종  $x$  의  $x$  를 씁니다.

주목. 역 요소는 연산  $\otimes$  에 대해 정의되며 반드시 다음을 의미하지는 않습니다 ◆

$$\frac{1}{x}$$

아벨 그룹

추가로  $\forall x, y \in G : x \otimes y = y \otimes x$  이면  $G = (G, \otimes)$  는 아벨 그룹 (가환성)입니다.

## 예 2.10(그룹)

연관된 연산이 있는 집합의 몇 가지 예를 살펴보고 집합이 그룹인지 여부를 확인합니다.

$\mathbb{N} \cup \{0\}$

- $(\mathbb{Z}, +)$  는 아벨 그룹입니다.
  - $(\mathbb{N}_0, +)$  는 그룹이 아닙니다.  $(\mathbb{N}_0, +)$  는 중성 요소 (0)를 가지고 있지만 역 요소는 없습니다.
  - $(\mathbb{Z}, \cdot)$  는 그룹이 아닙니다.  $(\mathbb{Z}, \cdot)$  에는 중성 요소 (1)가 포함되지만 모든  $z \in \mathbb{Z}, z \neq \pm 1$ 에 대한 역 요소는 없습니다.
  - $(\mathbb{R}, \cdot)$  은 0 이 역원소를 가지지 않기 때문에 군 이 아닙니다.
  - $(\mathbb{R} \setminus \{0\}, \cdot)$  는 아벨입니다.
  - $(\mathbb{R}_n, +), (\mathbb{Z}_n, +), n \in \mathbb{N}$  은  $+$  가 구성요소별로 정의된 경우 Abelian 입니다. 즉,  $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$ . (2.61)
- 그러면  $(x_1, \dots, x_n) \stackrel{+}{\mapsto} (-x_1, \dots, -x_n)$  은 역원소이고  $e = (0, \dots, 0)$  은 중성 원소이다.
- $(\mathbb{R}^m \times \mathbb{R}^n, +), m \times n$ -행렬 의 집합 은 Abelian입니다((2.61)에 정의된 대로 구성요소별 덧셈 포함).
  - $(\mathbb{R}^n \times \mathbb{R}^n, \cdot)$ , 즉 (2.13)에 정의된 행렬 곱셈이 있는  $n \times n$  행렬 의 집합을 자세히 살펴보겠습니다.

– 폐쇄와 연관성은 행렬 곱셈의 정의에서 직접 따릅니다.

– 중성 요소: 단위 행렬  $I_n$  은 다음을 갖는 중성 요소입니다.  
 $(\mathbb{R}^n \times \mathbb{R}^n, \cdot)$  의 행렬 곱셈 " $\cdot$ " 에 대해.



- 역 요소: 역 요소가 존재하는 경우 ( $A$ 는 정규),  $A \in \mathbb{R}^{n \times n}$ 의 역 요소 및 정확히  $A^{-1}$ 이다  
 경우  $(\mathbb{R}^{n \times n}, \cdot)$ 는 일반 선형 그룹이라고 하는 그룹입니다.

정의 2.8 (일반 선형 그룹). 정규(가역) 행렬  $A \in \mathbb{R}^{n \times n}$ 의 집합은 (2.13)에 정의된 행렬 곱셈에 대한 그룹이며 일반 선형 그룹  $GL(n, \mathbb{R})$ 이라고 합니다. 그러나 행렬 곱셈은 가환성이 아니므로 일반 선형 그룹은 Abelian 그룹이 아닙니다.

### 2.4.2 벡터 공간

그룹에 대해 논의할 때 집합  $G$ 와  $G$ 에 대한 내부 연산, 즉  $G$ 의 요소에 대해서만 연산을 수행하는 매핑  $G \times G \rightarrow G$ 를 살펴보았습니다. 다음에서는 내부 연산 + 곱셈에 추가하여 집합을 고려할 것입니다. 벡터  $x \in G$  by 는 외부 연산  $\cdot$ , 스칼라  $\lambda \in \mathbb{R}$ 도 포함합니다. 내부 연산은 덧셈의 형태로, 외부 연산은 스칼라 곱셈의 형태로 생각할 수 있습니다. 내부/외부 곱셈과 관련이 없습니다.

정의 2.9 (벡터 공간). 실수 벡터 공간  $V = (V, +, \cdot)$ 는 벡터 공간이 두 개의 연산을 갖는 집합  $V$ 입니다.

$$+ : V \times V \rightarrow V \quad (2.62)$$

$$\cdot : \mathbb{R} \times V \rightarrow V \quad (2.63)$$

여기

1.  $(V, +)$ 는 아벨군 2. 분포도: 1.

$$\forall \lambda \in \mathbb{R}, x, y \in V : \lambda \cdot (x + y)$$

$$= \lambda \cdot x + \lambda \cdot y \quad 2. \forall \lambda, \psi \in \mathbb{R}, x \in V : (\lambda + \psi) x = \lambda x + \psi x$$

3. 결합성(외부 연산):  $\forall \lambda, \psi \in \mathbb{R}, x \in V : \lambda \cdot (\psi \cdot x) = (\lambda \psi) \cdot x$  4. 외부 연산에 대한

중성 요소:  $\forall x \in V : 1 \cdot x = x$

요소  $x \in V$ 를 벡터라고 합니다.  $(V, +)$ 의 중성 요소는 0 벡터  $0 = [0, \dots, 0]^T$ , 내부 연산  $+$ 를 벡터 벡터 덧셈이라고 합니다. 요소  $\lambda \in \mathbb{R}$ 를 스칼라라고 하며 외부 연산 스칼라  $\cdot$ 는 스칼라를 곱한 것입니다. 스칼라 곱은 다른 값을 곱한 것이므로 섹션 3.2에서 설명합니다.

스칼라

주목. "벡터 곱셈"  $ab$ ,  $a, b \in \mathbb{R}^n$  적으로는  $c = ab$ 와  $c_j = a_j b_j$ , 정의되어 있지 않습니다. 이론가 되도록 요소별 곱셈을 정의할 수 있습니다. 이 "배열 곱셈"은 많은 프로그래밍 언어에 공통적이지만 행렬 곱셈에 대한 표준 규칙을 사용하면 수학적으로 제한된 의미를 갖습니다. 벡터를  $n \times 1$  행렬로 처리

(우리가 일반적으로 하는 방식), (2.13)에 정의된 행렬 곱셈을 사용할 수 있습니다. 그러나 벡터의 차원이 일치하지 않습니다. 벡터에 대한 다음 곱셈만 정의됩니다.  $a \cdot b \in \mathbb{R}^{n \times n}$  (외적),  $a \cdot b \in \mathbb{R}$  (내적/스칼라/내적). ♦

외부 제품

- $V = \mathbb{R}^n$ ,  $n \in \mathbb{N}$ 은 다음과 같이 정의된 연산이 있는 벡터 공간입니다. - 더하기:  $x+y = (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1+y_1, \dots, x_n+y_n)$  모든  $x, y \in \mathbb{R}^n$  - 스칼라 곱하기:  $\lambda x = \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$

모든  $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$ ,

- $V = \mathbb{R}^{m \times n}$ ,  $m, n \in \mathbb{N}$ 은 다음을 갖는 벡터 공간입니다.

$$a_{11} + b_{11} \quad \dots \quad 1n + b_{1n}$$

$$a_{m1} + b_{m1} \quad a_{mn} + b_{mn}$$

$$\lambda a_{11} \cdot \dots \cdot \lambda a_{1n}$$

$$\lambda a_{m1} \cdot \dots \cdot \lambda a_{mn}$$

■

$$, \mathbb{R}^{n \times 1},$$

열 벡터

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

행 벡터

&gt;

바꾸어 놓다

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.

## 2.4 벡터 공간

39

## 2.4.3 벡터 부분 공간 다음에

서 벡터 부분 공간을 소개합니다. 직관적으로, 그것들은 우리가 이 부분 공간 내의 요소들에 대해 벡터 공간 연산을 수행할 때, 우리는 결코 그것을 떠나지 않을 속성을 가진 원래 벡터 공간에 포함된 세트입니다. 이런 의미에서 그들은 "닫힌" 것입니다. 벡터 부분공간은 기계 학습의 핵심 아이디어입니다. 예를 들어, 10장에서는 차원 축소를 위해 벡터 부분 공간을 사용하는 방법을 보여줍니다.

정의 2.10 (벡터 부분공간).  $V = (V, +, \cdot)$  를 벡터 공간이라고 하고  $U \subseteq V, U \neq \emptyset$ . 그런 다음  $U = (U, +, \cdot)$  는  $U$  가 벡터 공간 연산  $+$  선형 부분 공간을 갖는 벡터 공간 이고  $\cdot U \times U$  및  $R \times U$  로 제한되는 경우  $V$  의 벡터 부분 공간 (또는 벡터 부분 공간 선형 부분 공간) 이라고 합니다.  $U \subseteq V$  를 써서  $V$  의 부분공간  $U$  를 나타냅니다.

$U \subseteq V$  및  $V$  가 벡터 공간이면  $U$  는 모든  $x \in V$ , 특히 모든  $x \in U \subseteq V$  에 대해 유지하기 때문에 자연스럽게  $V$  로부터 직접 많은 속성을 상속 합니다. 여기에는 Abelian 그룹 속성, distributivity, 연관성 및 중립 요소.  $(U, +, \cdot)$  가  $V$  의 부분 공간 인지 여부를 결정하려면 여전히 표시해야 합니다.

1.  $U \neq \emptyset$ , 특히:  $0 \in U$  2.  $U$  폐쇄:

a. 외부 연산에 관하여:  $\forall \lambda \in R \forall x \in U: \lambda x \in U$ . b. 내부 연산과 관련하여:  $\forall x, y \in U: x + y \in U$ .

## 예제 2.12(벡터 부분공간)

몇 가지 예를 살펴보겠습니다.

- 모든 벡터 공간  $V$  에 대해 그림 2.6, 사소한 부분 공간은  $V$  자체와  $\{0\}$ 입니다.
- 2.6의 유일한 예  $D$  는  $R^2$  의 부분 공간입니다 (일반적인 내부/외부 연산 포함).  $A$  와  $C$  에서 클로저 속성이 위반되었습니다.  $B$  는  $0$  을 포함하지 않습니다.
- 선형 방정식  $Ax = 0$  의 동종 시스템의 솔루션 세트  $n$  개의 미지수  $x = [x_1, \dots, x_n]^T$   $> R^n$  의 부분공간입니다.
- 선형 방정식  $Ax = b, b \neq 0$  의 비균일 시스템의 해는  $R^n$  의 부분공간이 아닙니다.
- 임의의 많은 부분 공간의 교집합은 부분 공간 자체입니다.

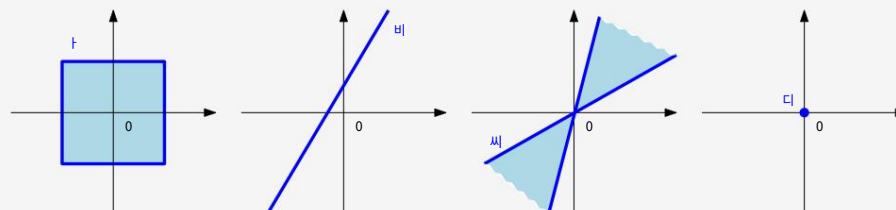


그림 2.6  $R^2$  의 모든 부분집합이 부분공간인 것은 아닙니다.  $A$  와  $C$  에서 클로저 속성이 위반되었습니다.  $B$  는  $0$  을 포함하지 않습니다.  $D$  만 부분공간입니다.

주목. 모든 부분공간  $U \subseteq (\mathbb{R}^n, +, \cdot)$  는 호모의 해 공간  $\diamond$   
 $x \in \mathbb{R}^n$  에 대해 선형 연립방정식  $Ax = 0$ .

## 2.5 선형 독립

다음에서는 벡터로 무엇을 할 수 있는지 자세히 살펴해보겠습니다.  
 (벡터 공간의 요소). 특히, 벡터를 함께 추가할 수 있습니다.  
 스칼라로 곱합니다. 폐쇄 속성은 우리가  
 같은 벡터 공간에서 다른 벡터로 끝냅니다. 찾을 수 있습니다  
 벡터의 모든 벡터를 표현할 수 있는 벡터 세트  
 함께 추가하고 크기를 조정하여 공간을 확보합니다. 이 벡터 세트는  
 2.6.1 절 에서 논의할 것입니다. 우리가 거기에 도착하기 전에,  
 선형 결합과 선형 결합의 개념을 도입해야 합니다.  
 독립.

정의 2.11 (선형 조합). 벡터 공간  $V$  와  $a$ 를 고려하십시오.  
 유한한 수의 벡터  $x_1, \dots, x_k \in V$ . 그런 다음 형식의 모든  $v \in V$

$$v = \lambda_1 x_1 + \dots + \lambda_k x_k = \sum_{i=1}^k \lambda_i x_i \in V \quad (2.65)$$

$\lambda_1, \dots, \lambda_k \in \mathbb{R}$  은 벡터  $x_1, \dots, x_k$ .

0 벡터 는  $k \text{ vec } 0x_i$ 의 선형 조합 이 항상 참이므로 항상 쓸 수 있습니다. 다음에서,  
 토르  $x_1, \dots, x_k$  의  $0 = \sum_{i=1}^k 0x_i$  이기 때문에  $x_k$   $\sum_{i=1}^k$   
 리는 벡터 세트의 중요하지 않은 선형 조합에 관심이 있습니다.

0, 즉 벡터  $x_1, \dots, x_k$ 의 선형 조합을 나타냅니다  $\dots, x_k$ , 모두가 아닌 경우  
 (2.65)의 계수  $\lambda_i$  는 0입니다.

정의 2.12 (선형(인)종속). 벡터 공간을 생각해보자  
 형  $\text{com } V$  가 있는 경우  $\dots, x_k \in V$ .  $0 = \sum_{i=1}^k \lambda_i x_i$  이기 때문에  $x_k$   $\sum_{i=1}^k$   
 $\lambda_i$  적어도 하나 의  $\lambda_i \neq 0$ 인  $\lambda_i x_i$ , 벡터

선형 종속 선형 독립

$x_1, \dots, x_k$  는 선형 종속적입니다. 사소한 솔루션만 존재한다면, 즉,  
 $\lambda_1 = \dots = \lambda_k = 0$  벡터  $x_1, \dots, x_k$  는 선형 독립입니다.

선형 독립은 선형에서 가장 중요한 개념 중 하나입니다.  
 대수학. 직관적으로 선형 독립 벡터 세트는 벡터로 구성됩니다.  
 중복성이 없는 것, 즉  
 세트, 우리는 무언가를 잃을 것입니다. 다음 섹션에서는  
 이 직감을 더 공식화하십시오.

### 예제 2.13(선형 종속 벡터)

지리적 예는 선형 독립의 개념을 명확히 하는 데 도움이 될 수 있습니다. 키갈리(르완다)가  
 어디에 있는지 설명하는 나이로비(케냐)의 한 사람  
 "우간다 캄팔라 에서 북서쪽으로 506km 이동한 다음 남서쪽 으로 374km 이  
 동하면 키갈리에 도착할 수 있습니다." 충분한 정보입니다

## 2.5 선형 독립

41

지리적 좌표 시스템이 2차원 벡터 공간으로 간주될 수 있기 때문에 키갈리의 위치를 설명하기 위해(고도 무시 그리고 지구의 곡면). 그 사람은 다음과 같이 덧붙일 수 있습니다. “약 751km입니다. 여기서 서쪽으로.” 이 마지막 진술은 사실이지만 다음과 같이 할 필요는 없습니다. 이전 정보가 주어지면 Kigali를 찾습니다(그림 2.7 참조). 이 예에서 "506km Northwest" 벡터(파란색)와 "374 km Southwest" 벡터(보라색)는 선형 독립입니다. 이것은 의미 남서 벡터는 북서 벡터로 설명할 수 없으며 그 반대도 마찬가지입니다. 그러나 세 번째 "751km West" 벡터(검은색)는 다른 두 벡터의 선형 결합은 벡터 집합을 선형 종속적으로 만듭니다. 동일하게 "서쪽 751km" 및 "374km Southwest"는 선형 결합하여 "506km Northwest"를 얻을 수 있습니다.



그림 2.7  
지리적 예  
(조  
근사치  
기본 방향)  
선형적으로  
종속 벡터  
안에  
2차원  
공간(비행기).

주목. 다음 속성은 벡터가  
선형 독립:

- $k$  벡터는 선형 종속 또는 선형 독립입니다. 거기 세 번째 옵션이 아닙니다.
- 벡터  $x_1, \dots, x_k$  가 0 이면 선형 종속적입니다. 두 벡터가 동일한 경우에도 마찬가지입니다.
- 벡터  $\{x_1, \dots, x_k : x_i \neq 0, i = 1, \dots, k\}$ ,  $k > 2$ , 선형 (적어도) 그들 중 하나가 선형 조합인 경우에만 종속 다른 사람의. 특히, 한 벡터가 다른 벡터의 배수인 경우, 즉,  $x_i = \lambda x_j$ ,  $\lambda \in \mathbb{R}$  다음 집합  $\{x_1, \dots, x_k : x_i \neq 0, i = 1, \dots, k\}$  선형 종속적입니다.
- 벡터  $x_1, \dots, x_k \in V$  는 선형 독립은 가우스 제거를 사용하는 것입니다. 모든 벡터를 열로 작성 행렬  $A$  의 행렬이 다음 과 같을 때까지 가우스 소거를 수행 행 계층 형식(여기서 축소 행 계층 형식은 필요하지 않음):

- 피벗 열은 왼쪽에 있는 벡터와 선형으로 독립적인 벡터를 나타냅니다. 행렬이 만들어질 때 벡터의 순서가 있음에 유의하십시오.
- non-pivot 열은 왼쪽에 있는 피벗 열의 선형 조합으로 표현할 수 있습니다. 예를 들어, 행 사다리꼴 형식

$$\begin{array}{cccccc} 1 & 3 & 0 & 0 & 0 \\ & 2 & & & \end{array} \quad (2.66)$$

첫 번째와 세 번째 열이 피벗 열임을 알려줍니다. 두 번째 열은 첫 번째 열의 세 배이기 때문에 피벗이 아닌 열입니다.

모든 열 벡터는 모든 열이 피벗 열인 경우에만 선형 독립적입니다. 피벗이 아닌 열이 하나 이상 있는 경우 열(따라서 해당 벡터)은 선형 종속적입니다.



예제 2.14 다음과 같은 R4를 고려하십시오.

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \quad x_3 = \begin{pmatrix} -1 \\ -2 \\ 1 \\ 1 \end{pmatrix}. \quad (2.67)$$

선형 종속 여부를 확인하기 위해 일반적인 접근 방식을 따르고

$$\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = \lambda_1 \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ -2 \\ 1 \\ 1 \end{pmatrix} = 0 \quad (2.68)$$

$\lambda_1, \dots, \lambda_3$ . 벡터  $x_i$   $i = 1, 2, 3$ 을 행렬의 열로 작성하고 피벗 열을 식별할 때까지 기본 행 연산을 적용합니다.

$$\begin{array}{cccc} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{array} \quad \dots \quad \begin{array}{ccc} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}. \quad (2.69)$$

여기에서 행렬의 모든 열은 피벗 열입니다. 따라서 사소하지 않은 솔루션은 없으며 방정식 시스템을 풀기 위해  $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0$ 이 필요합니다. 따라서 벡터  $x_1, x_2, x_3$ 은 선형 독립입니다.

## 2.5 선형 독립

43

주목.  $k$ 개의 선형 독립 벡터  $b_1, \dots, b_k$  및  $m$  선형 조합

$$\begin{aligned} x_1 &= \sum_{i=1}^k \lambda_{i1} b_i, \\ &\vdots \\ x_m &= \sum_{i=1}^k \lambda_{im} b_i \end{aligned} \quad (2.70)$$

$B = [b_1, \dots, b_k]$  열이 선형 독립 벡터  $b_1, \dots, b_k$ , 우리는 쓸 수 있습니다

$$x_j = B \lambda_j, \lambda_j = \begin{pmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{kj} \end{pmatrix}, j = 1, \dots, m. \quad (2.71)$$

더 컴팩트 한 형태로.

$x_1, \dots, x_m$  은 선형 독립입니다. 이를 위해  $\psi_j x_j = 0$ 입니다.  
 목적, 우리는 테스트의 일반적인 접근 방식을 따릅니다.  $\psi_j$  With  $j=1$   
 (2.71), 우리는

$$\sum_{j=1}^m \psi_j x_j = \sum_{j=1}^m \psi_j B \lambda_j = B \sum_{j=1}^m \psi_j \lambda_j. \quad (2.72)$$

이것은  $\{x_1, \dots, x_m\}$  은 열 벡터가  $\{\lambda_1, \dots, \lambda_m\}$  은 선형 독립입니다.



주목. 벡터 공간에서  $V$  는  $m$  ,  $k$  벡터 의  $m$  선형 조합  $x_1, \dots, x_k$  ◆  
 $> k$  인 경우 선형 종속적 입니다.

## Example 2.15

선형 독립 벡터  $b_1, b_2, b_3, b_4 \in \mathbb{R}^n$  및

$$\begin{aligned} x_1 &= b_1 - 2b_2 + b_3 - b_4 & x_2 &= -4b_1 - 2b_2 + 4b_4 \\ x_3 &= 2b_1 + 3b_2 - b_3 - 3b_4 & x_4 &= 17b_1 - 10b_2 + b_3 + b_4 \end{aligned} \quad (2.73)$$

벡터  $x_1, \dots, x_4 \in \mathbb{R}^n$  선형 독립? 이 질문에 답하기 위해 열 벡터가

$$\begin{pmatrix} -4 & 2 & 17 \\ 1 & -2 & -10 \\ 2 & 3 & -1 & 11 \\ -1 & 4 & -3 & 1 \end{pmatrix} \quad (2.74)$$

$$A = \begin{pmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -14 & -3 & 1 & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & \end{pmatrix}$$

해당하는 선형 방정식 시스템은 쉽게 풀 수 없습니다. 마지막 열은 피벗 열이 아니며  $x_4 = -7x_1 - 15x_2 - 18x_3$ 입니다.

$x_1$ 의 선형 조합, ...,  $x_3$ .

## 2.6 기준 및 순위

### 2.6.1 집합과 기저 생성

$(V, +, \cdot)$  및 벡터 집합  $A = \{x_1, \dots, x_k\} \subseteq V$ . 모든 벡터  $v \in V$ 는  $x_1, \dots, x_k, A$ 는  $a$

생성 세트  
기간

최소한의

기초

$A \sim (V$ 에 걸쳐 있는  $A \subseteq V$ .  $V$ 의 모든 선형 독립 생성 집합

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.



$V = (V, +, \cdot)$  를 벡터 공간이라고 하고  $B \subseteq V$ ,  $B \neq \emptyset$ 이라고 합니다. 그러면 다음 명제들이 동일합니다.

- $B$  는  $V$  의 기저입니다 .
- $B$  는 최소 생성 집합입니다.
- $B$  는  $V$  의 최대 선형 독립 벡터 집합입니다 . 즉, 이 집합에 다른 벡터를 추가하면 선형 종속이 됩니다.
- 모든 벡터  $x \in V$  는  $B$ 의 벡터의 선형 결합이며 모든 선형 결합은 고유합니다. 즉,

$$x = \sum_{i=1}^n \lambda_i b_i = \sum_{i=1}^n \psi_i b_i \quad (2.77)$$

및  $\lambda_i, \psi_i \in \mathbb{R}$ ,  $b_i \in B$ ,  $\lambda_i = \psi_i$ ,  $i = 1, \dots, n$ .

기초는 최소한의  
생성 집합 및 최대 선형 독립  
집합  
벡터.

### 예 2.16

- $\mathbb{R}^3$  에서, 표준/표준 기저 는

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.78)$$

표준 기저

- $\mathbb{R}^3$  의 다른 염기 는

$$B_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0.5 & 1.8 & -2.2 \\ 0.8 & 0.3 & -1.3 \\ 0.4 & 0.3 & 3.5 \end{pmatrix}. \quad (2.79)$$

- 세트

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & 1 \\ 3 & 0 & 0 \\ 4 & 2 & -4 \end{pmatrix} \quad (2.80)$$

선형 독립적이지만  $\mathbb{R}^4$  의 생성 집합(기저 없음)이 아닙니다 . 예를 들어, 벡터  $[1, 0, 0, 0]^T$  은  $A$  에 있는 요소의 선형 조합으로 얻을 수 없습니다 .

주목. 모든 벡터 공간  $V$  는 기저  $B$  를 가지고 있습니다 . 앞의 예는 벡터 공간  $V$  의 기저가 많을 수 있음을 보여줍니다 . 즉, 고유 기저가 없습니다. 그러나 모든 염기는 고유한 기저 벡터를 갖습니다.

◆ 기저 벡터

유한 차원 벡터 공간  $V$  만 고려합니다 . 이 경우,  $V$  의 차원 은  $V$  의 기저 벡터의 수입니다. , 그리고 우리는  $\dim(V)$  을 씁니다. 차수  $U \subseteq V$  가  $V$  의 부분공간인 경우 , 그런 다음  $\dim(U) \leq \dim(V)$  및  $\dim(U) =$

$\dim(V) = U = V$  인 경우에만. 직관적으로, 벡터 공간의 차원은 이 벡터 공간에서 독립적인 방향의 수로 생각할 수 있습니다.

차원  
벡터 공간은  
기저의 수에 해당합니다.  
벡터.

요소의 집합입니다. 예를 들어, 벡터 공간  $V = \text{span}[1\text{차원이지만 기본 벡터는 두 개의 요소 } 0]$  은 1를 가지고 있습니다. ♦

주목. 부분공간  $U$  의 기저  $= \text{span}[x_1, \dots, x_m] \subseteq \mathbb{R}^n$  은 다음 단계를 실행하여 찾을 수 있습니다.

♦

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 2 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 3 \\ -4 \\ 5 \\ -3 \end{pmatrix}, \quad x_4 = \begin{pmatrix} -1 \\ 8 \\ -6 \\ 1 \end{pmatrix} \quad 5 \in \mathbb{R},$$

4

$$\sum_{i=1}^4 \lambda_i x_i = 0, \quad \lambda_i \in \mathbb{R},$$

 $x_1, x_2, x_3, x_4 =$ 

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ -4 & 8 & -6 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & -2 \\ & & & & & & & 1 \end{pmatrix}$$

피벗 열은 어떤 벡터 집합이 선형으로 독립적인지 나타내므로 행-사각형 형식에서  $x_1, x_2, x_4$  가 선형 독립임을 알 수 있습니다(선형 방정식 시스템  $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_4 x_4 = 0$  은  $\lambda_1 = \lambda_2 = \lambda_4 = 0$ ). 따라서  $\{x_1, x_2, x_4\}$  는  $U$  의 기저가 됩니다.

### 2.6.2 순위

행렬  $A \in \mathbb{R}^{m \times n}$  의 선형 독립 열의 수는 선형 독립 행의 수와 같으며  $A$  의 순위 순위라고 하며  $\text{rk}(A)$  로 표시됩니다.

주목. 행렬의 순위에는 몇 가지 중요한 속성이 있습니다.

- $\text{rk}(A) = \text{rk}(A^T)$ , 즉 열 순위는 행 순위와 같습니다.
- $\mathbb{R}^{m \times n}$  의 열은  $\dim(U) = \text{rk}(A)$  인 부분 공간  $U \subseteq \mathbb{R}^n$  에 걸쳐 있습니다. 나중에 이 부분 공간을 이미지 또는 범위라고 부를 것입니다.  $U$  의 기저는 다음과 같습니다. 피벗 열을 식별하기 위해  $A$  에 가우스 소거법을 적용하여 찾았습니다.
- $A \in \mathbb{R}^{m \times n}$  의 행은  $\dim(W) = \text{rk}(A)$  인 부분공간  $W \subseteq \mathbb{R}^m$  에 걸쳐 있습니다.  $W$  의 기저는 가우스 소거법을 적용하여 찾을 수 있습니다.
- 모든  $A \in \mathbb{R}^{n \times n}$  에 대해  $\text{rk}(A) = n$  인 경우에만  $A$  가 규칙적(가역적) 임을 유지합니다.
- 모든  $A \in \mathbb{R}^{m \times n}$  및 모든  $b \in \mathbb{R}^m$  에 대해 선형 방정식 시스템  $Ax = b$  는  $\text{rk}(A) = \text{rk}(A|b)$  인 경우에만 풀릴 수 있으며, 여기서  $A|b$  는 증강 시스템을 나타냅니다.
- $A \in \mathbb{R}^{m \times n}$  에 대해  $Ax = 0$  에 대한 해의 부분공간은 차원  $n - \text{rk}(A)$  를 갖습니다. 나중에 이 부분 공간을 커널 또는 널 커널 널 공간이라고 부를 것입니다.
- 행렬  $A \in \mathbb{R}^{m \times n}$  은 순위가 동일한 차원의 행렬에 대해 가능한 가장 큰 전체 순위 순위와 같으면 전체 순위를 갖습니다. 이는 전체 순위 행렬의 순위가 행과 열의 수 중 작은 수임을 의미합니다(즉,  $\text{rk}(A) = \min(m, n)$ ). 순위가 없는 행렬이 전체 순위를 갖지 않으면 순위가 부족한 행렬이라고 합니다.



#### 예 2.18(순위)

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

$A$  는  $\text{rk}(A) = 2$  가 되도록 선형으로 독립된 두 개의 행/열을 갖습니다.

$$\blacksquare A = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 5 & 0 & 0 \end{pmatrix}$$

순위를 결정하기 위해 가우스 소거법을 사용합니다.

$$\begin{pmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.84)$$

여기서  $\text{rk}(A) = 2$  인 선형 독립 행과 열의 수가 2임을 알 수 있습니다.

## 2.7 선형 매핑

다음에서는 좌표 개념을 정의할 수 있도록 구조를 유지하는 벡터 공간에 대한 매핑을 연구합니다.

이 장의 시작 부분에서 벡터는 함께 더하고 스칼라로 곱할 수 있는 객체이며 결과 객체는 여전히 벡터라고 말했습니다. 매핑을 적용할 때 이 속성을 유지하기를 원합니다. 두 개의 실수 벡터 공간  $V, W$ 를 고려합니다. 매핑  $\Phi: V \rightarrow W$ 는 다음과 같은 경우 벡터 공간의 구조를 유지합니다.

$$\Phi(x + y) = \Phi(x) + \Phi(y) \quad (2.85)$$

$$\Phi(\lambda x) = \lambda \Phi(x) \quad (2.86)$$

모든  $x, y \in V$  및  $\lambda \in \mathbb{R}$ 에 대해 다음 정의로 요약할 수 있습니다.

선형 매핑 벡터  
공간 동형

정의 2.15 (선형 매핑). 벡터 공간  $V, W$ 의 경우 매핑  $\Phi: V \rightarrow W$ 를 선형 매핑 (또는 벡터 공간 동형/ 선형 변환)이라고 합니다.

선의  
변환

$$\forall x, y \in V \forall \lambda, \psi \in \mathbb{R} : \Phi(\lambda x + \psi y) = \lambda \Phi(x) + \psi \Phi(y). \quad (2.87)$$

선형 매핑을 행렬로 나타낼 수 있음이 밝혀졌습니다(섹션 2.7.1). 벡터 집합을 행렬의 열로 수집할 수도 있음을 상기하십시오. 행렬로 작업할 때 행렬이 나타내는 것이 무엇인지 염두에 두어야 합니다. 선형 매핑 또는 벡터 모음입니다. 선형 매핑에 대한 자세한 내용은 4장에서 볼 것입니다. 계속하기 전에 특수 매핑을 간단히 소개하겠습니다.

주사  
형용사

정의 2.16 (형사, 형용사, 전사). 매핑  $\Phi: V \rightarrow W$ 를 고려하십시오. 여기서  $V, W$ 는 임의의 집합일 수 있습니다. 그런 다음  $\Phi$ 가 호출됩니다.

- $\forall x, y \in V$ 인 경우 형용사:  $\Phi(x) = \Phi(y) \Rightarrow x = y$ .
- $\Phi(V) = W$ 인 경우 객관식입니다.
- 분사형 및 형용사인 경우 전단사.

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.

## 2.7 선형 매핑

49

$\Phi$  가 전수적 이면  $W$  의 모든 요소는  $\Phi$  를 사용하여  $V$  에서 "도달"할 수 있습니다. 전단사  $\Phi$  는 "취소"될 수 있습니다. 즉,  $\Psi \circ \Phi(x) = x$  가 되도록 매핑  $\Psi: W \rightarrow V$  가 존재합니다. 이 매핑  $\Psi$  는  $\Phi$  의 역함수라고 하며 일반적으로  $\Phi^{-1}$  로 표시됩니다.

이러한 정의를 통해 벡터 공간  $V$  와  $W$  사이의 선형 매핑에 대한 다음과 같은 특별한 경우를 소개합니다.

- 동형사상:  $\Phi: V \rightarrow W$  선형 및 전단사 내형사상:  $\Phi: V \rightarrow V$  선형
- 항동형사상:  $\Phi: V \rightarrow V$  선형 및 전단사 우리는  $V$  에서  $\text{id}_V: V \rightarrow V$  를 정의합니다.
- $\text{id}_V$  ,  $x \mapsto x$  는 ID 매핑 또는 ID 매핑 으로 사용

동형  
내형성  
자기형성

정체성  
자기형성

실시예 2.19(동형)

매핑  $\Phi: \mathbb{R}^2 \rightarrow \mathbb{C}$ ,  $\Phi(x) = x_1 + ix_2$  는 동형입니다.

$$\begin{aligned} \Phi(x_1 + y_1 + ix_2 + iy_2) &= (x_1 + y_1) + i(x_2 + y_2) = x_1 + ix_2 + y_1 + iy_2 \\ &= \Phi(x_1 + ix_2) + \Phi(y_1 + iy_2) \\ \Phi(\lambda(x_1 + ix_2)) &= \lambda(x_1 + ix_2) = \lambda\Phi(x_1 + ix_2) = \lambda\Phi(x) \end{aligned} \quad (2.88)$$

이것은 또한 복소수가  $\mathbb{R}^2$ 에서 튜플로 표시될 수 있는 이유를 정당화합니다.  $\mathbb{R}^2$  에 있는 튜플의 요소별 덧셈을 해당 덧셈이 있는 복소수 집합으로 변환하는 전단사 선형 매핑이 있습니다. 우리는 선형성을 보여 주었을 뿐 전단사는 보여주지 않았습니다.

정리 2.17 (Axler(2015)의 정리 3.59). 유한 차원 벡터 공간  $V$  와  $W$  는  $\dim(V) = \dim(W)$  인 경우에 만 동형입니다.

정리 2.17은 동일한 차원의 두 벡터 공간 사이에 선형 전단사 매핑이 존재한다고 말합니다. 직관적으로 이것은 동일한 차원의 벡터 공간이 손실 없이 서로 변환될 수 있기 때문에 일종의 동일한 것임을 의미합니다.

정리 2.17은 또한  $\mathbb{R}^m \times \mathbb{R}^n$  ( $m \times n$  행렬의 벡터 공간)과  $\mathbb{R}^{mn}$  (길이가  $mn$ 인 벡터의 벡터 공간) 을 차원이  $mn$ 이고 선형이 존재하기 때문에 동일 하게 취급하는 정당성을 제공합니다. 하나를 다른 것으로 변환하는 이중 형용사 매핑.

주목. 벡터 공간  $V, W, X$ 를 고려하십시오. 그런 다음:

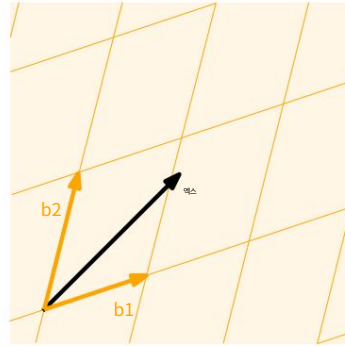
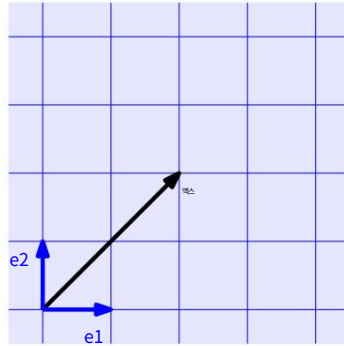
- 선형 매핑  $\Phi: V \rightarrow W$  및  $\Psi: W \rightarrow X$ 의 경우 매핑  $\Psi \circ \Phi: V \rightarrow X$  도 선형입니다.
- $\Phi: V \rightarrow W$  가 동형이면  $\Phi$  isomorphism 도 마찬가지입니다.  $\Phi^{-1}: W \rightarrow V$  는 이성질체

50

선형 대수학

그림 2.8 두 개의 다른 좌표  
두 가지 기저 세트로 정의된 시스템  
벡터. 벡터  $x$ 는 다른

동등 어구  
선택한 좌표계에 따른 표현.



$$B = (b_1, \dots, b_n)$$

주문 기준

따라서 여기에서 일부 부분을 요약합니다.  $B = (b_1, \dots, b_n)$ 은 순서가 있는 기저,  $B = \{b_1, \dots, b_n\}$ 은 (순서가 없는) 기저이고  $B = [b_1, \dots, b_n]$ 은 열이 벡터  $b_1, \dots, b_n$ 인 행렬.

$$x = \alpha_1 b_1 + \dots + \alpha_n b_n$$

동등 어구

$$A^{-1} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

좌표 벡터  
동등 어구  
대표

"기계 학습을 위한 수학" 초안(2022-01-11). 피드백: <https://mml-book.com>.

## 2.7 선형 매핑

51

기준 좌표계를 효과적으로 정의합니다. 우리는 표준 기저 벡터  $e_1, e_2$ 에 의해 확장되는 2차원의 데카르트 좌표계에 익숙합니다. 이 좌표계에서 벡터  $x \in \mathbb{R}^2$ 는  $e_1$ 과  $e_2$ 를 선형적으로 결합하여  $x$ 를 얻는 방법을 알려주는 표현을 가지고 있습니다. 그러나  $\mathbb{R}^2$ 의 모든 기저는 유효한 좌표계를 정의하고 이전의 동일한 벡터  $x$ 는  $(b_1, b_2)$  기저에서 다른 좌표 표현을 가질 수 있습니다. 그림 2.8에서 표준기저  $(e_1, e_2)$ 에 대한  $x$ 의 좌표는  $[2, 2]^T$ 이다. 그러나 기저  $(b_1, b_2)$ 에 대해서는 동일한 벡터  $x$ 가  $[1.09, 0.72]^T$ , 즉  $x = 1.09b_1 + 0.72b_2$ 로 표현된다. 다음 섹션에서 이 표현을 얻는 방법을 알아보겠습니다.

## Example 2.20

좌표가  $[2, 3]^T$ 인 기하 벡터  $x \in \mathbb{R}^2$ 를  $\mathbb{R}^2$ 의 표준 기저  $(e_1, e_2)$ 에 대해 살펴보겠습니다. 이것은  $x = 2e_1 + 3e_2$ 라고 쓸 수 있음을 의미합니다. 그러나 이 벡터를 나타내기 위해 표준 기저를 선택할 필요는 없습니다. 기저 벡터  $b_1 = [1, -1]^T, b_2 = [1, 1]^T$ 를 사용하여 동일한 벡터를 나타내면  $(b_1, b_2)$  (그림 2.9 참조).

12

기초.

$$x = 2e_1 + 3e_2$$

$$x = 2b_1 + 5b_2$$

주목.  $n$ 차원 벡터 공간  $V$  및  $V$ 의 정렬된 기저  $B$ 에 대해  $\dots, n$ 은 선형이며(정리 2.17로 인해 행렬이 됨) 여기서  $(e_i)_{i=1}^n$ 은  $\mathbb{R}^n$ 의 표준 기저입니다.

이제 행렬과 행렬 사이를 명시적으로 연결할 준비가 되었습니다. 유한 차원 벡터 공간 간의 선형 매핑.

정의 2.19 (변환 매트릭스). 대응하는(정렬된) 밑  $B = (b_1, \dots, b_n)$  및  $C = (c_1, \dots, c_m)$ 인 벡터 공간  $V, W$ 를 고려하십시오. 또한 선형 매핑  $\Phi: V \rightarrow W$ 를 고려합니다.  $j \in \{1, \dots, n\}$ ,

$$\Phi(b_j) = a_{1j}c_1 + \dots + a_{mj}c_m = X_m \quad \text{아이언스} \quad (2.92)$$

나=1

는  $C$ 에 대한  $\Phi(b_j)$ 의 고유한 표현입니다. 그런 다음  $m \times n$ -행렬  $A_\Phi$ 를 호출합니다.

$$A_\Phi(i, j) = a_{ij}, \quad (2.93)$$

$\Phi$ 의 변환 행렬 ( $V$  변환의 정렬된 베이스  $B$  및  $W$ 의  $C$ 에 대해).

행렬

$W$ 의 정렬된 기저  $C$ 에 대한  $\Phi(b_j)$ 의 좌표는  $A_\Phi$ 의  $j$ 번째 열입니다. (유한 차원) 벡터 공간  $V, W$ , 정렬된 밑  $B, C$  및 선형 매핑  $\Phi: V \rightarrow W$

변환 행렬  $A\Phi$ .  $x$  가  $B$  에 대한  $x \in V$  의 좌표 벡터이고  $y$  가  $C$  에 대한  $y = \Phi(x) \in W$  의 좌표 벡터인 경우

$$y = A\Phi x. \quad (2.94)$$

이는 변환 행렬을 사용하여  $V$  의 정렬된 기준에 대한 좌표를  $W$  의 정렬된 기준에 대한 좌표로 매핑할 수 있음을 의미합니다.

예제 2.21(변환 행렬)

동형사상  $\Phi: V \rightarrow W$  및  $V$  의 정렬된 염기  $B = (b_1, \dots, b_3)$  및  $W$  의  $C = (c_1, \dots, c_4)$  를 고려 합니다.

$$\begin{aligned} \Phi(b_1) &= c_1 - c_2 + 3c_3 - c_4 \\ \Phi(b_2) &= 2c_1 + c_2 + 7c_3 + 2c_4 \\ \Phi(b_3) &= 3c_2 + c_3 + 4c_4 \end{aligned} \quad (2.95)$$

변환 행렬  $A\Phi$  는  $\Phi(b_k) = \sum_{i=1}^4 a_{ik} c_i$  을 만족합니다.  $k = 1$  에 대한  $\Phi(b_1)$  을  $a_{11}c_1 + a_{21}c_2 + a_{31}c_3 + a_{41}c_4$  로 나타내면 다음과 같이 주어진다.

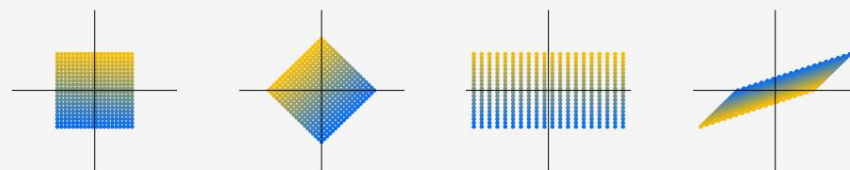
$$A\Phi = [\alpha_1, \alpha_2, \alpha_3] = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}, \quad (2.96)$$

여기서  $\alpha_j, j = 1, 2, 3$  은  $\Phi(b_j)$  의 좌표 벡터입니다.  $C$  에게.

예제 2.22(벡터의 선형 변환)

그림 2.10 선형 변환의 세 가지 예

(a)에서 점으로 표시된 벡터; (b)  $45^\circ$  회전 (c) 수평 좌표를 2만큼 늘린다. (d) 반사, 회전 및 스트레칭의 조합.



(a) 원본 데이터.

(b)  $45^\circ$  회전. (c) 수평축을 따라 늘린다.

(d) 일반 선형 매핑.

변환 행렬을 사용하여  $\mathbb{R}^2$  의 벡터 세트에 대한 세 가지 선형 변환을 고려 합니다.

$$A_1 = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 3 \\ -1 & -1 \end{bmatrix}. \quad (2.97)$$



## 2.7 선형 매핑

53

그림 2.10은 벡터 세트의 선형 변환에 대한 세 가지 예를 보여줍니다. 그림 2.10(a)는  $R^2$  의 400개 벡터를 보여줍니다. , 각각이 대표되는 해당  $(x_1, x_2)$  좌표에서 점으로. 벡터의 범위는 정사각형입니다. (2.97)에서 행렬  $A_1$  을 사용하여 선형 변환할 때 이 벡터들 각각에 대해 그림 2.10(b)에서 회전된 정사각형을 얻습니다. 만약 우리가  $A_2$ 로 표시되는 선형 매핑을 적용하면 다음에서 직사각형을 얻습니다. 그림 2.10(c)는 각  $x_1$  좌표를 2만큼 늘린다. 그림 2.10(d) 선형 변환할 때 그림 2.10(a)의 원래 제곱을 보여줍니다. 반사, 회전 및 늘이기의 조합인  $A_3$ 를 사용 합니다.

## 2.7.2 베이스스 변경

다음에서는 변환 행렬이 어떻게 이루어지는지 자세히 살펴보겠습니다. 선형 매핑  $\Phi: V \rightarrow W$  는  $V$  의 밑을 변경하면 변경 되고  $W$ . 두 개의 정렬된 염기를 고려하십시오.

$$B = (b_1, \dots, b_n), B \sim = (\sim b_1, \dots, \sim b_n) \quad (2.98)$$

$V$  와 2 개의 정렬된 염기

$$C = (c_1, \dots, c_m), C \sim = (c \sim 1, \dots, c \sim m) \quad (2.99)$$

of  $W$ . 또한,  $A_\Phi \in R_m \times n$  은 선형의 변환 행렬이다 .

매핑  $\Phi$  : 베이스  $B$  및  $C$ 에 대한  $V \rightarrow W$  및  $A \sim$

$$\Phi \in R_m \times n$$

는  $B \sim$  및  $C \sim$ 에 대한 해당 변환 매핑입니다 .

다음에서 우리는  $A$  와  $A \sim$  가 어떻게 관련되어 있는지 조사할 것입니다.

$A_\Phi$  를  $A \sim$ 로 변환할 수 있는지 여부

피 우리가 기초를 수행하기로 선택한 경우

$B, C$  에서  $B$ 로 변경 , 비교. 우리.

는 효과적으로 서로 다른 좌표 표현을 얻습니다.

ID 매핑  $\text{id}_V$  . 그림 2.9의 맥락에서 이것은 다음을 의미합니다.

$(e_1, e_2)$  에 대한 좌표를 에 대한 좌표로 매핑

$(b_1, b_2)$  벡터  $x$  를 변경하지 않고 기저를 변경하고 이에 따라 벡터의 표현을 변경함으로써

이 새로운 기초와 관련하여 다음을 허용하는 특히 간단한 형식을 가질 수 있습니다.

간단한 계산을 위해. ♦

실시에 2.23(베이스스 변경)  
변환 행렬 고려

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad (2.100)$$

$R^2$  의 표준 기반과 관련하여 .

새로운 근거를 정의한다면

$$B = \left( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \right) \quad (2.101)$$

우리는 대각 변환 행렬을 얻습니다

$$\sim = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.102)$$

B 와 관련하여 A 보다 작업하기가 더 쉽습니다 .

다음에서는 한 기준에 대한 좌표 벡터를 다른 기준에 대한 좌표 벡터로 변환하는 매핑을 살펴보겠습니다. 먼저 주요 결과를 설명한 다음 설명을 제공합니다.

정리 2.20 (기저 변경). 선형 매핑의 경우  $\Phi: V \rightarrow W$ , 정렬된 염기

$$B = (b_1, \dots, b_n), B \sim = (\sim b_1, \dots, \sim b_n) \quad (2.103)$$

$V$  및  $_$

$$C = (c_1, \dots, c_m), C \sim = (\sim c_1, \dots, \sim c_m) \quad (2.104)$$

$W$ 의 변환 행렬  $A_\Phi$  및  $B$  와  $C$  에 대한  $\Phi$  의 변환 행렬  $A \sim$  기저  $B \sim$  및  $C \sim$ 에 대한 대응 변환 행렬  $A \sim$  는 다음 과 같이 주어집니다.

$$\mathbf{t}_{\sim} = T^{-1} A_\Phi S. \quad (2.105)$$

여기서,  $S \in \mathbb{R}^n \times \mathbb{R}^n$  은  $B \sim$  에 대한 좌표를  $B$ 에 대한 좌표에 매핑하는  $\text{id}_V$ 의 변환 행렬이고,  $T \in \mathbb{R}^m \times \mathbb{R}^m$  은  $C \sim$ 에 대한 좌표를 좌표에 매핑 하는  $\text{id}_W$ 의 변환 행렬이다.  $C$  와 관련하여

Proof following Drumm and Weil(2001), 우리는  $V$ 의 새로운 기저  $B \sim$ 의 벡터를  $B$ 의 기저 벡터의 선형 조합으로 쓸 수 있습니다.  
그런

$$\sim b_j = s_{1j}b_1 + \dots + s_{nj}b_n = \sum_{i=1}^n s_{ij}b_i, j = 1, \dots, n. \quad (2.106)$$

유사하게, 우리는  $W$ 의 새로운 기저 벡터  $C \sim$ 를  $C$ 의 기저 벡터의 선형 조합으로 씁니다 .

$$\sim c_k = t_{1k}c_1 + \dots + t_{mk}c_m = \sum_{l=1}^m t_{lk}c_l, k = 1, \dots, m. \quad (2.107)$$

우리는  $S = ((s_{ij})) \in \mathbb{R}^n \times \mathbb{R}^n$  을  $B \sim$ 에 대한 좌표를  $B$ 에 대한 좌표에 매핑하는 변환 행렬로 정의하고  $T = ((t_{lk})) \in \mathbb{R}^m \times \mathbb{R}^m$  을 좌표를 매핑하는 변환 행렬로 정의합니다.  $C \sim$ 에 대한 좌표는  $C$ 에 대한 좌표입니다 . 특히,  $S$ 의  $j$ 번째 열은  $B$ 에 대한  $\sim b_j$ 의 좌표 표현 이고

## 2.7 선형 매핑

55

T의 k번째 열은 C에 대한  $c \sim k$ 의 좌표 표현입니다. S와 T는 모두 규칙적입니다.

우리는 두 가지 관점에서  $\Phi(\sim bj)$ 를 살펴볼 것입니다. 먼저 매핑  $\Phi$ 를 적용하여 모든  $j = 1, \dots, N$

$$\Phi(\sim bj) = \sum_{k=1}^m a_{\sim k} c_{\sim k} \quad (2.107) = \sum_{k=1}^m a_{\sim k} \sum_{l=1}^n x_{kl} c_l = \sum_{l=1}^n \left( \sum_{k=1}^m a_{\sim k} x_{kl} \right) c_l \quad \text{엑스 엠} \quad (2.108)$$

여기서 우리는 먼저 새로운 기저 벡터  $c \sim k \in W$ 를 기저 벡터  $c_l \in W$ 의 선형 조합으로 표현한 다음 합산 순서를 바꿨습니다.

또는  $\sim bj \in V$ 를 다음의 선형 조합으로 표현하면  $b_j \in V$ , 우리는 도착

$$\Phi(\sim bj) \stackrel{(2.106)}{=} \Phi \sum_{n=1}^n x_{nj} b_n = \sum_{n=1}^n x_{nj} \Phi(b_n) \quad \text{알리클(2.109a)}$$

$$\sum_{n=1}^n x_{nj} \Phi(b_n) = \sum_{n=1}^n x_{nj} \sum_{l=1}^n s_{lj} \Phi(b_l) \quad (2.109b)$$

여기서 우리는  $\Phi$ 의 선형성을 이용했습니다. (2.108)과 (2.109b)를 비교하면 모든  $j = 1, \dots, n$  및  $l = 1, \dots, n$ 에 대해

$$\sum_{k=1}^m a_{\sim k} x_{kl} = \sum_{n=1}^n s_{lj} x_{nj} \quad \text{알리시} \quad (2.110)$$

따라서,

$$\Phi \sim bj = A \Phi S \in \mathbb{R}^{m \times n}, \quad (2.111)$$

그런

$$\Phi \sim bj = T^{-1} A \Phi S, \quad (2.112)$$

정리 2.20을 증명합니다.  $\square$

정리 2.20은  $V$  ( $B$ 가  $B \sim$ 로 대체 됨) 및  $W$  ( $C$ 가  $C \sim$ 로 대체 됨)의 기저 변화로 선형 매핑  $\Phi: V \rightarrow W$ 의 변환 행렬  $A \Phi$ 가 등가 행렬  $A$ 로 대체됨을 알려줍니다.  $\sim$

$\Phi$ 와 함께

$$\Phi \sim bj = T^{-1} A \Phi S. \quad (2.113)$$

그림 2.11은 이 관계를 보여줍니다. 동형  $\Phi: V \rightarrow W$ 와  $V$ 와  $C$ 의 정렬된 염기  $B, B \sim, W$ 의  $C \sim$ 를 고려합니다. 매핑  $\Phi C B$ 는  $\Phi$ 의 인스턴스화이며  $B$ 의 기저 벡터를 기저의 선형 조합에 매핑합니다.  $C$ 의 벡터. 정렬된 염기  $B, C$ 에 대해  $\Phi C B$ 의 변환 행렬  $A \Phi$ 를 알고 있다고 가정합니다.

그림 2.11

동형

 $\Phi: V \rightarrow W$  및정렬된 염기  $B, B \sim$   
 $V$  및  $C, W$ 의  $C \sim$   
(파란색으로 표시),  
우리는 표현할 수 있습니다 $\Phi C \sim B \sim$  매핑

기초에 대한 존중

 $B \sim C \sim$ 와 동등하게

의 조성

동형사상

 $\Phi C \sim B \sim =$  $\Xi C C \sim \circ \Phi C B \circ \Psi B B \sim$ 

에 관하여

의 기초

참자. 그만큼

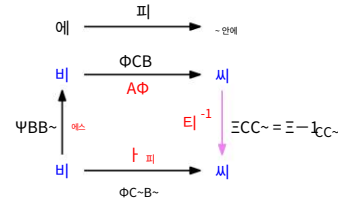
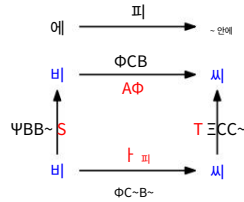
동

변환

매트릭스는 빨간색입니다.

벡터 공간

주문 기지

대응하는 변환 행렬  $A \sim$  $\Psi B B \sim$ 의 행렬 표현을 찾습니다. $\Phi$ 는 다음과 같습니다. 먼저 선형 매핑

"이전" 기준  $B$ 에 대해 ( $V$ 에서). 그런 다음  $\Phi C B$ 의 변환 행렬  $A \Phi: V \rightarrow W$ 를 사용하여 이 좌표를 좌표에 매핑합니다.

$W$ 의  $C$ 에 대해. 마지막으로 선형 매핑  $\Xi C C \sim$ 를 사용합니다.  $W \rightarrow W$   
 $C$ 에 대한 좌표에 대한 좌표에 매핑합니다.

씨. 따라서 선형 매핑  $\Phi C \sim B \sim$ 를 다음의 구성으로 표현할 수 있습니다.

"이전" 기반을 포함하는 선형 매핑:

$$\Phi C \sim B \sim = \Xi C C \sim \circ \Phi C B \circ \Psi B B \sim = \Xi C C \sim \circ \Phi C B \circ \Psi B B \sim. \quad (2.114)$$

구체적으로, 우리는  $\Psi B B \sim = \text{id}_V$  및  $\Xi C C \sim = \text{id}_W$ , 즉 ID 매핑을 사용합니다.

벡터를 자체에 매핑하지만 다른 기준에 대해 매핑합니다.

동등한

정의 2.21 (동등성). 두 행렬  $A, A \sim \in \mathbb{R}^m \times n$ 

동등하다

정규 행렬  $S \in \mathbb{R}^n \times n$  및  $T \in \mathbb{R}^m \times m$ 이 존재하는 경우  
 $A \sim = T^{-1}AS$ .

비슷한

정의 2.22 (유사성). 두 행렬  $A, A \sim \in \mathbb{R}^n \times n$ 

비슷한 경우 \_

$A \sim = S^{-1}AS$ 인 정규 행렬  $S \in \mathbb{R}^n \times n$ 이 존재합니다.

주목. 유사한 행렬은 항상 동일합니다. 그러나 등가 행렬이 반드시 유사하지는 않습니다.

◆

주목. 벡터 공간  $V, W, X$ 를 고려하십시오. 다음 설명에서

정리 2.17, 선형 매핑  $\Phi: V \rightarrow W$

및  $\Psi: W \rightarrow X$  매핑  $\Psi \circ \Phi: V \rightarrow X$ 도 선형입니다. 와 함께

대응하는 매핑의 변환 행렬  $A \Phi$  및  $A \Psi$ ,

전체 변환 행렬은  $A \Psi \circ \Phi = A \Psi A \Phi$ 입니다. ◆

이 말에 비추어 우리는 perspec에서 기본 변경 사항을 볼 수 있습니다.

선형 매핑 구성:

- $A \Phi$ 는 선형 매핑  $\Phi C B$ 의 변환 행렬입니다.  $V \rightarrow W$   
베이스  $B, C$ 에 대해
- $\Phi C \sim B \sim$ 의 변환 행렬:  $V \rightarrow W$   
베이스  $B$ 에 대해  $S$ 는 선형 매핑 씨.
- $\Psi B B \sim$ 의 변환 행렬입니다.  $V \rightarrow V$   
(automorphism)  $B \sim$ 를  $B$ 로 표현합니다. 일반적으로  $\Psi = \text{id}_V$ 는  
 $V$ 의 ID 매핑

## 2.7 선형 매핑

57

- $T$  는 선형 매핑  $\exists C \sim : W \rightarrow W$  (automorphism) 의 변환 행렬로  $C \sim$  를  $C$  로 표현합니다. 일반적으로  $\exists = \text{id}_W$  는  $W$  의 동일성 매핑입니다.

우리가 (비공식적으로) 변환을 밑수로만 기록하면  $A\Phi : B \rightarrow C$ ,  $A \sim \Phi : B \sim \rightarrow C \sim$ ,  $S : B \sim \rightarrow B$ ,  $T : C \sim \rightarrow C$  및  $\text{티}^{-1} : C \rightarrow C \sim$ , 및

$$B \sim \rightarrow C \sim = B \xrightarrow{\sim} B \xrightarrow{\sim} C \rightarrow C \sim \quad (2.115)$$

$$\text{티}^{-1} = T^{-1}A\Phi S. \quad (2.116)$$

(2.116)의 실행 순서는 오른쪽에서 벡터를 공급하기 때문에 오른쪽에서 왼쪽으로  $x \xrightarrow{7} Sx \xrightarrow{7} A\Phi(Sx) \xrightarrow{7} T^{-1}A\Phi(Sx) = A \sim \Phi x$ .

실시예 2.24(베이스스 변경)

변환 행렬이 다음과 같은 선형 매핑  $\Phi : R^3 \rightarrow R^4$  를 고려합니다.

$$A\Phi = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{pmatrix} \quad (2.117)$$

표준 베이스와 관련하여

$$B = \left( \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, C = \left( \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right). \quad (2.118)$$

변환 행렬  $A \sim$ 를 구합니다.

새로운 염기에 대한  $\Phi$  의

$$B \sim = \left( \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \right) \in R^3, \quad C \sim = \left( \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right). \quad (2.119)$$

그 다음에,

$$\text{에스} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad \text{티} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2.120)$$

여기서  $S$  의  $i$  번째 열은  $B$  의 기저 벡터에 대한  $\sim b_i$  의 좌표 표현입니다.  $B$  가 표준 기저이므로 좌표 표현을 쉽게 찾을 수 있습니다. 일반 기저  $B$  에 대해 다음 과 같은  $\lambda_i$  를 찾기 위해 선형 방정식 시스템을 풀어야 합니다.

P3.  $\lambda_i b_i = \sim b_j, j = 1, \dots, 3$ . 마찬가지로  $T$ 의  $j$ 번째 열은 좌표입니다.  
 $C$ 의 기저 벡터에 대한  $c \sim j$ 의 고유 표현.  
 따라서 우리는

$$T^{-1} \Phi S = \begin{matrix} & & 1 & -1 & -1 & & & & 3 & 2 & 1 \\ \frac{1}{2} & 1 & 1 & -1 & 1 & -1 & & & 0 & 4 & 2 \\ & -1 & 1 & & 1 & & 1 & & 10 & 8 & 4 \\ & & 0 & 0 & 0 & 2 & & & 1 & 6 & 3 \end{matrix} \quad (2.121a)$$

$$= \begin{matrix} -4 & -4 & -2 \\ 6 & 0 & 4 & 8 & 6 \\ & 0 & & & \\ & 4 & & & \\ 1 & & & & \end{matrix} \cdot \begin{matrix} 0 \\ 4 \\ \text{삼} \end{matrix} \quad (2.121b)$$

4장에서 우리는 기초 변화의 개념을 활용할 수 있을 것입니다.  
 엔도모피즘의 변환 행렬이 특히 단순한(대각선) 형태를 갖는 기준을 찾는 것. 10장에서 우리는  
 데이터 압축 문제를 살펴보고 편리한 기반을 찾습니다.  
 압축 손실을 최소화하면서 데이터를 투영할 수 있습니다.

### 2.7.3 이미지와 커널

선형 매핑의 이미지와 커널은 특정 중요한 속성을 가진 벡터 부분 공간입니다. 다음에서 우리는 그들을 특성화 할 것입니다  
 더 조심스럽게.

정의 2.23 (이미지 및 커널).

$\Phi: V \rightarrow W$ 의 경우 커널/널 공간을 정의 합니다.

$$\ker(\Phi) := \Phi^{-1}(0_W) = \{v \in V : \Phi(v) = 0_W\} \quad (2.122)$$

및 이미지/범위

$$\text{Im}(\Phi) := \Phi(V) = \{w \in W \mid \exists v \in V : \Phi(v) = w\}. \quad (2.123)$$

우리는 또한  $V$ 와  $W$ 를 각각  $\Phi$ 의 도메인 과 코 도메인이라고 부릅니다.

직관적으로 커널은  $\Phi$ 가 매핑 하는 벡터  $v \in V$ 의 집합입니다.  
 중립 요소  $0_W \in W$ . 이미지는 다음과 같은 벡터  $w \in W$ 의 집합입니다.  
 $V$ 의 모든 벡터에서  $\Phi$ 로 "도달"할 수 있습니다. 삽화가 주어진다  
 그림 2.12.

주목. 선형 매핑  $\Phi: V \rightarrow W$ 를 고려합니다. 여기서  $V, W$ 는 벡터입니다.  
 공백.

- 항상  $\Phi(0_V) = 0_W$  이므로  $0_V \in \ker(\Phi)$ 를 유지합니다. ~ 안에  
 특히, null space는 결코 비어 있지 않습니다.
- $\text{Im}(\Phi) \subseteq W$ 는  $W$ 의 부분 공간 이고  $\ker(\Phi) \subseteq V$ 는  $V$ 의 부분 공간 입니다.

## 2.7 선형 매핑

59

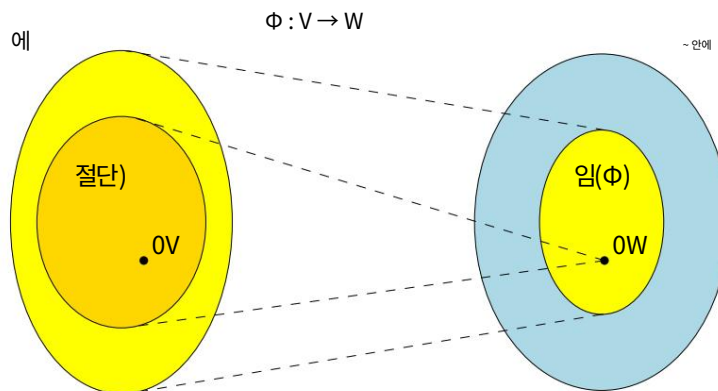


그림 2.12 선형 매핑의 커널과 이미지

 $\Phi: V \rightarrow W$ .

- $\Phi$ 는  $\ker(\Phi) = \{0\}$  인 경우에만 주입형(일대일)입니다.



비고 (Null Space 및 Column Space).  $A \in \mathbb{R}^m \times n$  및 선형 매핑  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto Ax$  를 고려합니다.

- $A = [a_1, \dots, a_n]$ , 여기서  $a_i$  는  $A$ 의 열이며 다음을 얻습니다.

$$\begin{aligned} \text{Im}(\Phi) &= \{Ax : x \in \mathbb{R}^n\} = \left\{ \sum_{i=1}^n x_i a_i : x_i \in \mathbb{R} \right\} \quad (2.124a) \\ &= \text{스팬}[a_1, \dots, a_n] \subseteq \mathbb{R}^m, \end{aligned}$$

- 
- 

- 
- 



$$4F: \mathbb{R}^2 \rightarrow \mathbb{R}^4, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ -x_1 \\ x_1 \\ 0 \end{pmatrix}$$





## 2.8 아핀 공간

61

## 2.8 아핀 공간

다음에서는 오프셋된 공간에 대해 자세히 살펴보겠습니다.  
원점, 즉 더 이상 벡터 부분공간이 아닌 공간. 게다가 우리는  
이러한 아핀 공간 간의 매핑 속성에 대해 간략하게 논의할 것입니다.  
선형 매핑과 유사합니다.

주목. 기계 학습 문헌에서 선형  
affine은 때때로 명확하지 않아 affine에 대한 참조를 찾을 수 있습니다.  
공간/매핑을 선형 공간/매핑으로. ♦

## 2.8.1 아핀 부분공간

정의 2.25 (아핀 부분공간).  $V$  를 벡터 공간,  $x_0 \in V$  라고 하고  
 $U \subseteq V$  a 부분공간. 그런 다음 하위 집합

$$\text{패} = x_0 + U := \{x_0 + u : u \in U\} \quad (2.130a)$$

$$= \{v \in V \mid \exists u \in U : v = x_0 + u\} \subseteq V \quad (2.130b)$$

$V$  의 아핀 부분공간 또는 선형 다양체 라고 합니다.  $U$  는 방향 또는 아핀 부분공간 이라고 합니다.  
방향 공간,  $x_0$  을 지지점 이라고 합니다. 12장에서는 다음을 참조합니다.  
초평면 과 같은 부분공간 .

선형 매니폴드  
방향  
방향 공간  
지지점  
초평면

아핀 부분 공간의 정의는  $x_0 \in U$  인 경우 0 을 제외합니다.  
따라서 아핀 부분공간은 (선형) 부분공간(벡터 부분공간)이 아닙니다.  
 $x_0 \in U$  에 대한  $V$ .

아핀 부분 공간의 예는  $\mathbb{R}^3$  의 점, 선 및 평면입니다. , 어느  
(필연적으로) 원점을 통과하지 마십시오.

주목. a의 두 개의 아핀 부분공간  $L = x_0 + U$  및  $L' = x'_0 + U'$ 를 고려하십시오 .  
벡터 공간  $V$  . 그러면,  $L \subseteq L' \iff U \subseteq U'$  이고  $x_0 - x'_0 \in U'$  인 경우에만 가능합니다 .

아핀 부분공간은 종종 매개변수 로 설명됩니다.  $k$  차원을 고려하십시오 If  $(b_1, \dots, b_k)$  는  
sional affine space  $L = x_0 + U$  of  $V$  .  
 $U$ , 모든 요소  $x \in L$  은 다음과 같이 고유하게 설명될 수 있습니다.

$$x = x_0 + \lambda_1 b_1 + \dots + \lambda_k b_k, \quad (2.131)$$

여기서  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ . 이 표현을 매개변수 방정식 매개변수 방정식 이라고 합니다.  
방향 벡터  $b_1, \dots, b_k$  및 매개변수  $\lambda_1, \dots, \lambda_k$ . ♦ 매개변수

## 예 2.26(아핀 부분공간)

- 1차원 아핀 부분공간을 선이라고 하며 선 으로 쓸 수 있습니다.  
 $y = x_0 + \lambda b_1$ , 여기서  $\lambda \in \mathbb{R}$  및  $U = \text{span}[b_1] \subseteq \mathbb{R}^n$  은  $\mathbb{R}^n$  의 1차원 부분공간입니다 . 이것은  
선이 지지점  $x_0$  과 방향을 정의 하는 벡터  $b_1$  에 의해 정의됨을 의미합니다. 그림 2.13 참조  
설명을 위해.

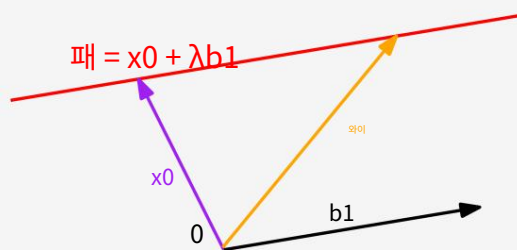
비행기

- $R^n$ 의 2차원 아핀 부분공간을 평면이라고 합니다. 평면에 대한 매개변수 방정식은  $y = x_0 + \lambda_1 b_1 + \lambda_2 b_2$ 입니다. 여기서  $\lambda_1, \lambda_2 \in R$  및  $U = \text{span}[b_1, b_2] \subseteq R^n$ 입니다. 이것은 평면이 지지점  $x_0$ 과 방향 공간에 걸쳐 있는 두 개의 선형 독립 벡터  $b_1, b_2$ 에 의해 정의됨을 의미합니다.

초평면

- $R^n$ 에서,  $(n-1)$ 차원의 아핀 부분 공간을 초평면이라고 합니다. 대응하는 매개변수 방정식은  $y = x_0 + \sum_{i=1}^{n-1} \lambda_i b_i$ 이  $(n-1)$ 차원 부분공간  $U = \text{span}[b_1, \dots, b_{n-1}]$ 에 걸쳐 있는  $x_0$ 와  $(n-1)$ 개의 선형 독립 벡터  $b_1, \dots, b_{n-1}$ 에 의해 정의됩니다.  $R^3$ 에서 평면은 초평면이기도 합니다.

그림 2.13 선은 아핀 부분 공간입니다. 선  $x_0 + \lambda b_1$ 의 벡터  $y$ 는 지지점  $x_0$  및 방향  $b_1$ 을 갖는 아핀 부분공간  $L$ 에 있습니다.



비고 (선형 방정식 및 아핀 부분 공간의 비균일 시스템).

$A \in R^{m \times n}$  및  $x \in R^m$ 에 대해 선형 방정식 시스템  $A\lambda = x$ 의 해는  $n - \text{rk}(A)$ 차원의  $R^n$ 의 빈 집합 또는 아핀 부분 공간입니다. 특히, 선형 방정식  $\lambda_1 b_1 + \dots + \lambda_n b_n = x$ , 여기서  $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$ 은  $R^n$ 의 초평면입니다.

$R^n$ 에서, 모든  $k$ 차원 아핀 부분 공간은  $\text{inhomogeneous}$ 의 솔루션입니다.

일차 연립방정식  $Ax = b$ , 여기서  $A \in R^{m \times n}$  및  $\text{rk}(A) = n - k$ . 동차 방정식 시스템,  $Ax = 0$ 의 경우 솔루션은 벡터 부분 공간이었고, 지지점  $x_0 = 0$ 인 특수 아핀 공간으로도 생각할 수 있습니다. ♦

### 2.8.2 아핀 매핑

섹션 2.7에서 논의한 벡터 공간 간의 선형 매핑과 유사하게 두 개의 아핀 공간 간의 아핀 매핑을 정의할 수 있습니다.

선형 및 아핀 매핑은 밀접하게 관련되어 있습니다. 따라서 선형 매핑에서 이미 알고 있는 많은 속성(예: 선형 매핑의 구성이 선형 매핑임)은 아핀 매핑에도 적용됩니다.

정의 2.26 (아핀 매핑). 두 벡터 공간  $V, W$ 에 대해 선형

## 2.9 추가 읽기

63

mapping  $\Phi : V \rightarrow W$ , 그리고  $a \in W$ , mapping

$$\Phi : V \rightarrow W \quad (2.132)$$

$$x \mapsto a + \Phi(x) \quad (2.133)$$

$V$ 에서  $W$ 로의 아핀 매핑입니다. 벡터  $a$ 를 번역 아핀 매핑이라고 합니다.

$\Phi$ 의 벡터.

번역 벡터

- 모든 아핀 매핑  $\Phi : V \rightarrow W$ 는 선형의 구성이기도 합니다.  
매핑  $\Phi : V \rightarrow W$  및 번역  $\tau : W \rightarrow W$  in  $W$ , 그러한  
 $\Phi = \tau \circ \Phi$  매핑  $\Phi$  및  $\tau$ 는 고유하게 결정됩니다.
- 구성  $\Phi \circ \tau$  아핀 매핑의  $\Phi : V \rightarrow W$ ,  $\tau : W \rightarrow X$ 는  
아핀.
- 아핀 매핑은 기하학적 구조를 불변으로 유지합니다. 또한 차원과 평행도를 유지합니다.

## 2.9 추가 읽기

Strang(2003), Golan(2007), Axler(2015) 및 Liesen의 교과서를 포함하여 선형 대수학 학습을 위한 많은 리소스가 있습니다.

메르만(2015). 이 장의 소개에서 언급한 몇 가지 온라인 리소스도 있습니다. 여기서는 가우시안 소거만 다루었지만,

선형 방정식, 그리고 우리는 수치 선형 대수학 교과서를 다음과 같이 참조합니다.

Stoer와 Burlirsch(2002), Golub과 Van Loan(2012), Horn과

Johnson(2013)을 참조하십시오.

이 책에서 우리는 선형 대수학(예:

벡터, 행렬, 선형 독립, 기저) 및 관련 주제

벡터 공간의 기하학. 3장에서는 내부를 소개합니다.

규범을 유도하는 제품. 이러한 개념을 통해 각도를 정의할 수 있습니다.

길이와 거리, 직교 투영에 사용할 것입니다. 예측은 다음과 같은 많은 기계 학습 알고리즘에서 핵심적인 것으로 밝혀졌습니다.

선형 회귀 및 주성분 분석, 둘 다

각각 9장과 10장에서 다룹니다.

## 수업 과정

$$\text{무엇? } b := ab + a + b, \quad a, b \in \mathbb{R} \setminus \{-1\}$$

$$\text{삼 ? 엑스 ? } x = 15$$

$$\begin{aligned} \overline{k} &= \{x \in \mathbb{Z} \mid x - k = 0(\text{mod } n)\} \\ &= \{x \in \mathbb{Z} \mid \exists a \in \mathbb{Z}: (x - k = n \cdot a)\}. \end{aligned}$$

$$\text{아연} = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}$$

$$a \oplus b := \overline{a + b}$$

$$a \otimes b = \overline{a \times b},$$

$$\text{지} = \begin{matrix} 1xz \\ 01 \text{ 및} \\ 001 \end{matrix}$$

수업 과정

65

ㄱ.

$$\begin{array}{cc} 12 & 110 \\ 45 & 011 \\ 78 & 101 \end{array}$$

비.

$$\begin{array}{cc} 123 & 110 \\ 456 & 011 \\ 789 & 101 \end{array}$$

씨.

$$\begin{array}{cc} 110 & 123 \\ 011 & 456 \\ 101 & 789 \end{array}$$

디.

$$\begin{array}{ccc} & & 03 \\ 121 & 2 & 1-1 \\ 41-1-4 & & 21 \\ & & 52 \end{array}$$

392.

$$\begin{array}{cccc} 03 & & & \\ 1-1 & 12 & 1 & 2 \\ 21 & 41-1-4 & & \\ 52 & & & \end{array}$$

2.5 다음 비균일 선형 시스템  $Ax = b$ 의  $x$ 에서 모든 해의 집합  $S$ 를 구합니다. 여기서  $A$ 와  $b$ 는 다음과 같이 정의됩니다.

ㄱ.

$$A = \begin{array}{ccc} 1 & 1-1-1 & \\ 25-7-5 & & \\ 2-11 & & \\ 52-42 & & \end{array}, \quad b = \begin{array}{c} 1 \\ -2 \\ 4 \\ 6 \end{array}$$

비.

$$A = \begin{array}{ccc} 1-100 & 1 & \\ 1 & 10-30 & \\ 2-101-1 & & \\ -120-2-1 & & \end{array}, \quad b = \begin{array}{c} 3 \\ 6 \\ 5 \\ -1 \end{array}$$

2.6 가우스 소거법을 사용하여 불균일 방정식의 모든 해를 구합니다.  
시스템  $Ax = b$

$$A = \begin{array}{ccc} 010010 & & 2 \\ 000110 & & -1 \\ 010001 & & 1 \end{array}, \quad b = \begin{array}{c} 2 \\ -1 \\ 1 \end{array}$$



2.12  $R^4$ 의 두 부분 공간을 고려하십시오.

$$U_1 = \text{스팬} \begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 1 \\ -3 & 0 & -1 \\ 1 & -1 & 1 \end{bmatrix}, \quad U_2 = \text{스팬} \begin{bmatrix} -1 & 2 & -3 \\ -2 & -2 & 6 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}.$$

$U_1 \cap U_2$ 의 기저를 결정합니다.

2.13 두 개의 부분 공간  $U_1$ 과  $U_2$ 를 고려하십시오. 여기서  $U_1$ 은 동차 방정식 시스템  $A_1x = 0$ 의 솔루션 공간이고  $U_2$ 는 다음을 포함하는 동차 방정식 시스템  $A_2x = 0$ 의 솔루션 공간입니다.

$$A_1 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -2 & -1 \\ 2 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & -3 & 0 \\ 1 & 2 & 3 \\ 7 & -5 & 2 \\ 3 & -1 & 2 \end{bmatrix}.$$

†.  $U_1, U_2$ 의 차수를 결정합니다. 비.  $U_1$ 과  $U_2$ 의 밑을 결정 합니다. 씨.  $U_1 \cap U_2$ 의 기저를 결정합니다.

2.14 두 개의 부분 공간  $U_1$ 과  $U_2$ 를 고려하십시오. 여기서  $U_1$ 은  $A_1$ 과  $U_2$ 는  $A_2$ 의 열로 확장됩니다.

$$A_1 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -2 & -1 \\ 2 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & -3 & 0 \\ 1 & 2 & 3 \\ 7 & -5 & 2 \\ 3 & -1 & 2 \end{bmatrix}.$$

†.  $U_1, U_2$ 의 차수 결정 b.  $U_1$ 과  $U_2$ 의 기저를 결정 합니다. c.  $U_1 \cap U_2$ 의 기저 결정

2.15  $F = \{(x, y, z) \in R^3 \mid x+y-z=0\}$  및  $G = \{(a-b, a+b, a-3b) \mid a, b \in R\}$ . †.  $F$ 와  $G$ 가  $R^3$ 의 부분공간임을 보여 주십시오. 비. 기저 벡터에 의존하지 않고  $F \cap G$ 를 계산 합니다. 씨.  $F$ 에 대해 하나의 기저를 찾고  $G$ 에 대해 하나를 찾고 이전에 찾은 기저 벡터를 사용하여  $F \cap G$ 를 계산 하고 이전 질문으로 결과를 확인합니다.

2.16 다음 매핑은 선형입니까?

†.  $a, b \in R$ 이라고 하자.

$$1 \text{ 층: } \Phi: ([a, b]) \rightarrow R$$

$$f \mapsto \Phi(f) = \int_a^b f(x) dx,$$

1 여기서  $L([a, b])$ 는  $[a, b]$ 에 대한 적분 가능한 함수의 집합을 나타냅니다. 비.

$$\rightarrow C^0 \quad \Phi: C^0 \rightarrow C^0$$

$$f \mapsto \Phi(f) = f,$$

여기서  $k > 1$ ,  $C^k$  테일러 함 연속적으로 다른  $k$  번 집합을 나타냅니다. 0은 연속 함수 집합을 나타냅니다.





