선형대수학

정 주 희

2023년 3월 6일

차 례

제 1 장 $\mathbb{R}^n$ **1**
1.1 벡터의 정의와 기본 연산 1
1.2 벡터의 사영 6
1.3 평면 유클리드 기하학 8
1.4 벡터의 외적 14
1.5 직선과 평면의 벡터 방정식 16

제 2 장 벡터공간 **21**
2.1 부분공간, 기저, 차원 21
2.2 사영과 직교기저 32
2.3 직교여공간 37

제 3 장 행렬 **39**
3.1 행렬곱 39
3.2 고교수학에서 다루었던 행렬 44
3.3 선형연립방정식과 가우스 소거 47
3.4 선형사상 54
3.5 좌표벡터와 표상행렬 62
3.6 디터미넌트 69

제 4 장 토픽들 **75**
4.1 사영과 근사 75
4.1.1 선형방정식의 근사해 75
4.1.2 회귀(Regression) 76
4.2 좌표축의 회전 78
4.3 직교행렬 80
4.4 행렬의 고윳값, 고유벡터, 대각화 80
4.5 행렬의 분해 80

찾아보기 **81**

1

$\mathbb{R}^n$

1.1 벡터의 정의와 기본 연산

벡터란 크기와 아울러 방향을 가진 것이라고 흔히 정의한다. 이 정의는 벡터를 쉽게 소개하려는 소기의 목적은 달성할 수 있겠지만, 엄격히 말하자면 옳지 않다. 예를 들어 3차원 공간에서의 회전은 크기(회전 각도)와 방향(회전축)을 가지고 있지만 두 개의 회전 사이에 교환법칙이 성립하지 않기 때문에 벡터가 아닌 것이다. 더 정확히 말하자면 "벡터공간의 공리"를 만족하는, "덧셈"이라고 부르는, 물리·기하학적인 의미를 가진, "회전들 간의 2항연산"을 정의할 수 없다. 벡터공간의 공리에 대해서는 나중에 다시 공부하게 될 것이다.

벡터란 "좌표계의 변환에 의하여 변하지 않는 물리량"이라고 말하는 것은 상당히 좋은 정의이다. 하지만 이 정의는 많은 설명이 부가되어야만 그 의미가 명확해 질 것이다.

벡터란 본질적으로 물리적 개념이기 때문에 수학적으로는, 즉 집합론적으로는 완벽하게 정의할 수 없다. 수학적으로 할 수 있는 최선의 벡터의 정의는 어떤 특정한 공리들을 만족하는 수학적 구조체를 벡터공간이라 하고 그 원소를 벡터라고 부르는 것이다. 이 장에서는 일단 벡터공간의 가장 대표적이고 중요한 예인 $\mathbb{R}^n$을 공부하기로 한다. 일반적인 벡터공간은 제2장에서 공부하게 될 것이다.

정의 1.1 고정된 각각의 자연수 n에 대해서 다음과 같이 정의한다.

$\mathbb{R}^n$의 원소를 벡터(*vector*)라고 부르고, $\mathbb{R}$의 원소, 즉 실수를 스칼라(*scalar*)라고 부르기로 한다. $\mathbb{R}^n$은 벡터공간(*vector space*)이라고 한다.

벡터의 연산(*operation*)에는 2항연산 $+$, 1항연산 $-$ 및 스칼라곱(*scalar multiplication*)의 3가지가 있으며 아래와 같이 정의된다.

$\vec{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$, $\vec{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$, $a \in \mathbb{R}$일 때

$$\vec{u} + \vec{v} \stackrel{\text{def}}{=} (u_1 + v_1, \ldots, u_n + v_n),$$

$$-\vec{u} \stackrel{\text{def}}{=} (-u_1, \ldots, -u_n), \quad \vec{u} - \vec{v} \stackrel{\text{def}}{=} \vec{u} + (-\vec{v})$$
$$a\vec{u} \stackrel{\text{def}}{=} (au_1, \ldots, au_n).$$

$\vec{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ 일 때 각 $i = 1, \ldots, n$ 에 대하여 u_i 를 $\vec{u}$ 의 i 번째 성분(*component*)이라고 한다.

모든 성분이 0인 벡터, 즉 $(0, \ldots, 0)$ 형태의 벡터를 영벡터(*zero vector*)라고 부르고 $\vec{0}$로 나타낸다. ⊣

팩트 1.2 벡터 연산의 기본 성질을 아래에 나열해 보았다. 증명은 대단히 쉬우므로 생략한다.

(1) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ (교환법칙)

(2) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ (결합법칙)

(3) $\vec{u} + \vec{0} = \vec{u}$ (항등원)

(4) $\vec{u} - \vec{u} = \vec{0}$ (역원)

(5) $1\vec{u} = \vec{u}$

(6) $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ (스칼라곱 분배법칙1)

(7) $(a + b)\vec{u} = a\vec{u} + b\vec{u}$ (스칼라곱 분배법칙2)

(8) $(ab)\vec{u} = a(b\vec{u})$ (스칼라곱 결합법칙) ⊣

[정의 1.1]의 기하적 의미를 생각해 보자. $n = 2$ 일 때와 $n = 3$ 인 경우만 생각하면 충분할 것이다. 4차원 이상의 공간은 기하적 직관으로 파악하기가 어렵다.

이 정의에 따르면 벡터는 좌표공간 내의 점(point)과 똑같아 보인다. 통상 우리는 점을 P 와 같은 영문 대문자로 나타내고 원점 O 에서 점 P 에 이르는 화살표를 벡터 $\vec{p}$, 혹은 $\overrightarrow{OP}$ 로 나타낸다. [정의 1.1]은 원점을 출발점으로 하는 화살표는 그것의 끝점을 가지고 나타내기로 정한 것에 다름 아니다.

점은 본디 기하적 개념인데, $\mathbb{R}^2$ 에서는 우리는 이것을 실수의 순서쌍 (a, b) 로 표상(represent)한다. 원점에서 출발하여 점 (a, b) 에 이르는 벡터는 점과는 구별되는 개념이지만 우리는 이것의 표상으로 역시 (a, b) 를 사용한다. 즉 점과 벡터는 관련이 있기는 하지만 동일하지는 않은 개념인데, 이 둘은 표상을 실수의 순서쌍이라는 형식으로 공유하는 것이다.[†]

출발점이 원점이 아닌 화살표는 어떻게 나타내는가? 우리는 그 화살표의 출발점을 원점으로 옮겨서 생각한다. 그러니까 점 $P = (a_1, b_1)$ 에서 출발하여 점 $Q = (a_2, b_2)$

[†] 복소수도 $a + bi = (a, b)$ 로 본다면 역시 표상을 공유한다고 말할 수 있다.

에서 끝나는 화살표는 순서쌍 $(a_2 - a_1, b_2 - b_1)$으로 나타낸다. 길이와 방향이 같은 화살표는 동일한 것으로 본다고 생각하면 된다.

오른쪽 그림에 몇 개의 벡터의 예를 보였다.

$\vec{a} = (2, 1) = \overrightarrow{PQ}$

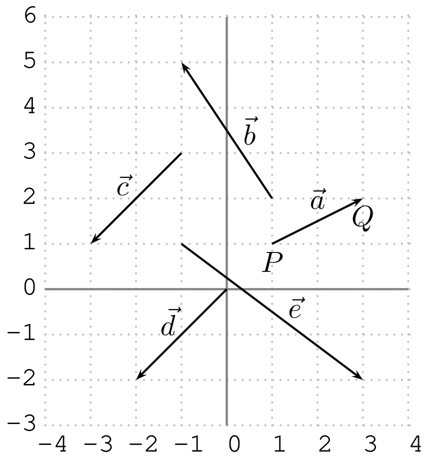
$\vec{b} = (-2, 3)$

$\vec{c} = (-2, -2) = \vec{d}$

$\vec{e} = (4, -3)$

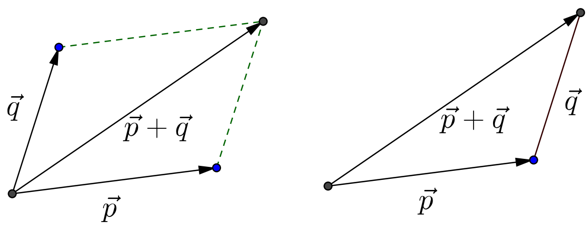
$\|\vec{a}\| = \sqrt{5}$ ($norm$ $\|\cdot\|$의 정의는 조금 후에 나온다.)

$\|\vec{e}\| = 5$



벡터합은 아래에 보인 [그림 1.1]로 이해하면 쉽다.

그림 1.1: 벡터의 합



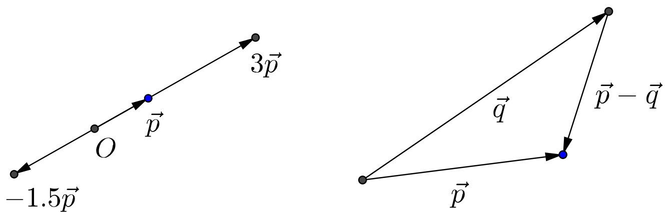
두 벡터 $\vec{p}$와 $\vec{q}$의 합 $\vec{p} + \vec{q}$는, 이 벡터들의 출발점을 일치시켰을 때, 이 벡터들이 나타내는 화살표들을 인접한 변으로 가지는 평행사변형에서 대각선 벡터에 대응된다. 이 사실은 간단한 유클리드 기하에 의하여 확인할 수 있다. 벡터의 합을 기하학적으로 해석하는 또 하나의 방법은 다음과 같다: $\vec{p}$의 끝점을 $\vec{q}$의 출발점에 일치시키면 $\vec{p} + \vec{q}$는 $\vec{p}$의 출발점에서 시작하여 $\vec{q}$의 끝점에 이르는 벡터가 된다.

벡터 $\vec{p}$, 스칼라 a가 주어졌을 때 이들의 스칼라곱 $a\vec{p}$는 $\vec{p}$의 방향은 그대로 두고 길이만 a배로 바꿔주는 것이다. 단 a가 음수일 때는 길이를 $|a|$배 늘려 주고, 방향을 반대로 하면 된다. Scalar를 읽을 때 스칼라라고 하는 것보다 사전에 나온 발음기호를 따라, 영어권 사람들이 발음하는 대로 스케일러(scale + er)라고 하는 것이 옳다고 본다. 스케일은 크기라는 뜻이니까 스케일러는 "크기를 정해주는 것"의 의미를 가지게 된다.

벡터의 빼기로 얻은 $\vec{p} - \vec{q}$는 여기에 $\vec{q}$를 더하면 $\vec{p}$가 되어야 한다. 따라서 $\vec{p} - \vec{q}$는 $\vec{q}$의 끝점에서 출발하여 $\vec{p}$의 끝점으로 가는 화살표가 나타내는 벡터가 된다.

스칼라곱과 벡터의 빼기는 [그림 1.2]에 나타내었다.

그림 1.2: 스칼라곱과 벡터의 빼기



정의 1.3 벡터간의 곱

$$\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

는 *dot product*, *scalar product* 혹은 내적(*inner product*)이라고 부르며 다음과 같이 정의된다.[†]

$$\vec{u}\cdot\vec{v} = u_1v_1 + \cdots + u_nv_n$$

벡터의 크기(*norm*), 또는 길이(*length*)는 내적을 이용하여 아래와 같이 정의된다.

$$\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$$

이 정의가 타당함은 피타고라스 정리로부터 쉽게 알 수 있다.

길이가 1인 벡터를 단위벡터(*unit vector*)라고 부른다. ⊣

벡터의 놈(norm)은 피타고라스 정리에 의하여 벡터의 길이를 의미하게 된다. 그리고 $\|\vec{u}\| = 0$는 $\vec{u} = \vec{0}$일 필요충분조건이다. 내적은 교환법칙과 분배법칙을 따른다. 그리고 스칼라곱과 내적을 혼합한 연산에 대해서는 $a(\vec{u}\cdot\vec{v}) = (a\vec{u})\cdot\vec{v} = \vec{u}\cdot(a\vec{v})$가 성립한다.[‡]

연습문제 1.4 다음을 만족하는 벡터 $\vec{a}, \vec{b} \in \mathbb{R}^2$를 구하여라.

(1) $\vec{u}\cdot\vec{v} > 0$

(2) $\vec{u}\cdot\vec{v} < 0$

(3) $\vec{u}\cdot\vec{v} = 0$

(4) $\|\vec{u}\| = 1$이고 $\vec{u}$의 두 성분은 모두 유리수. ⊣

[†] Scalar multiplication과 scalar product라는 두 용어는 생김새가 비슷하지만 의미는 대단히 다른 것이므로 혼동하지 않도록 주의해야 한다. 우리말로는 전자를 스칼라곱, 후자를 내적이라고 부르기로 한다.

[‡] 우리는 스칼라로 $\mathbb{R}$만을 사용하므로 이러하다. 스칼라로 $\mathbb{C}$를 사용하는 벡터간에는 연산법칙이 조금 다르다.

다음은 벡터의 길이와 내적에 대한 중요한 정리이다.

<u>정리</u> *1.5* $\vec{u}, \vec{v} \in \mathbb{R}^n$ 일 때

$$|\vec{u} \cdot \vec{v}| \le \|\vec{u}\| \, \|\vec{v}\|, \quad (\textit{Schwarz inequality}) \tag{1.1}$$

$$|\|\vec{u}\| - \|\vec{v}\|| \le \|\vec{u} + \vec{v}\| \le \|\vec{u}\| + \|\vec{v}\|, \quad (\textit{Triangle inequality}) \tag{1.2}$$

$$\|\vec{u} + \vec{v}\| = \|\vec{u} - \vec{v}\| \;\Leftrightarrow\; \vec{u} \cdot \vec{v} = 0. \tag{1.3}$$

(증명). (1.1)만 증명하고 (1.2)와 (1.3)은 연습문제로 남겨두겠다. 먼저 $\vec{e}$가 단위벡터인 경우에 $|\vec{u} \cdot \vec{e}| \le \|\vec{u}\| \, \|\vec{e}\| = \|\vec{u}\|$를 증명하겠다. $c = \vec{u} \cdot \vec{e}$로 놓는다. 그러면

$$0 \le (\vec{u} - c\vec{e}) \cdot (\vec{u} - c\vec{e}) = \|\vec{u}\|^2 - 2c(\vec{u} \cdot \vec{e}) + c^2 = \|\vec{u}\|^2 - c^2$$

이므로

$$\|\vec{u}\| \ge |c| = |\vec{u} \cdot \vec{e}| \tag{1.4}$$

를 얻는다. 일반적인 경우에는 (1.4)의 양변에 $\|\vec{v}\|$를 곱하고 $\vec{e} = \frac{1}{\|\vec{v}\|}\vec{v}$로 놓으면 원하는 부등식을 얻게 된다. (물론 $\vec{v} = \vec{0}$일 때는 별도로 처리해야 한다.) □

<u>단평</u> 1.6 Schwarz inequality는 코시-슈바르츠 부등식(*Cauchy-Schwarz inequality*)이라고도 하며 ① $\|\vec{u}\| \, \|\vec{v}\| \ge |\vec{u} \cdot \vec{v}|$에 더하여 ② 등호는 $\vec{u} = k\vec{v}$, 또는 $\vec{v} = k\vec{u}$인 $k \in \mathbb{R}$가 있을 때면이 성립한다는 사실을 포함하여 말한다.

여기서 ①을 [정리 1.5]에서의 증명과 다른 방법으로 증명하고 이어서 ②를 증명해 보기로 한다.

$\vec{v} = \vec{0}$인 경우에는 ①은 당연히 성립하고 ②는 $k = 0$로 놓아 성립한다. 따라서 아래의 논의에서는 $\vec{v} \ne \vec{0}$를 가정한다.

$\vec{u} = (u_1, \ldots, u_n)$, $\vec{v} = (v_1, \ldots, v_n)$로 놓고

$$(u_1^2 + \cdots + u_n^2)(v_1^2 + \cdots + v_n^2) \ge (u_1 v_1 + \cdots u_n v_n)^2 \tag{1.5}$$

을 보여야 한다. 모든 실수 t에 대하여 다음의 부등식이 성립함에 주목한다.

$$(u_1 - v_1 t)^2 + \cdots + (u_n - v_n t)^2 \ge 0 \tag{1.6}$$

t에 대한 2차식으로 정리하면

$$\|\vec{v}\|^2 t^2 - 2(\vec{u} \cdot \vec{v})t + \|\vec{u}\|^2 \ge 0 \tag{1.7}$$

모든 t에 대하여 위의 부등식이 성립한다는 것은 (2차항의 계수가 0이 아니므로) 판별식이 0 이하라는 것을 함의한다. 즉,

$$(\vec{u} \cdot \vec{v})^2 - \|\vec{u}\|^2 \, \|\vec{v}\|^2 \le 0$$

이 성립하며 이것이 바로 우리가 보이려던 부등식 ①이다.

이제 ②를 보이기 위하여 $\|\vec{u}\|^2\|\vec{v}\|^2 = (\vec{u}\cdot\vec{v})^2$을 가정하자. 그러면 판별식이 0이며 따라서 (1.7)의 좌변의 값, 즉 (1.6)의 좌변의 값이 0이 되는 $t \overset{\text{def}}{=} k$가 존재한다. 제곱들의 합이 0이라는 것은 각 항이 0이라는 것을 의미하므로 $u_1 = kv_1, \ldots, u_n = kv_n$이 성립한다. 즉 $\vec{u} = k\vec{v}$이다.

역방향의 증명은 아주 쉬우므로 생략한다. □

<u>연습문제</u> 1.7 [정리 1.5]의 식 (1.2)와 (1.3)을 각각 증명하여라.

(1.2)의 중변을 $\|\vec{u}-\vec{v}\|$로 바꾸어 얻은 부등식을 증명해 보라.

(1.2)에서 등호는 언제 성립하는가? ⊣

1.2 벡터의 사영

$\vec{0}$ 아닌 두 벡터 $\vec{u}$와 $\vec{v}$가 주어졌을 때, 이들을 인접한 변으로 가지는 평행사변형의 두 대각선 $\overrightarrow{BC} = \vec{u}-\vec{v}$와 $\overrightarrow{AD} = \vec{u}+\vec{v}$를 생각해 보자. 이 두 대각선들의 길이 $\|\vec{u}-\vec{v}\|$와 $\|\vec{u}+\vec{v}\|$는 이 평행사변형이 직사각형일 때, 그리고 이때만 같다는 사실을 유크리드 논증기하를 써서 증명할 수 있다. 즉, $\|\vec{u}-\vec{v}\| = \|\vec{u}+\vec{v}\|$는 $\vec{u}$와 $\vec{v}$가 서로 수직일 필요충분조건이 된다. 그러므로 (1.3)에 의하여 다음의 사실이 성립한다.

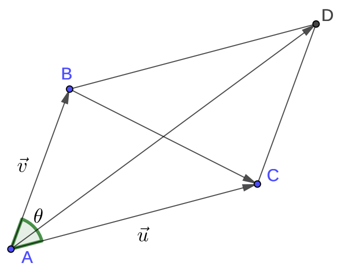


그림 1.3: 평행사변형의 두 대각선

<u>팩트</u> 1.8 $\mathbb{R}^2$와 $\mathbb{R}^3$에서는, 두 벡터는 그들의 내적이 0일때 서로 수직, 즉 *orthogonal*(*or normal or perpendicular*) 하다. ⊣

일반적인 $\mathbb{R}^n$에서 두 벡터가 수직하다는 것은 둘의 내적이 0인 것을 의미하는 것으로 정의한다. 이 정의에 따르면 영벡터는 모든 벡터와 수직임에 유의하라.

<u>정리</u> *1.9* 두 벡터 $\vec{p}$와 $\vec{q}$가 이루는 각을 θ라고 하면

$\vec{p}\cdot\vec{q} = \|\vec{p}\|\,\|\vec{q}\|\cos\theta$ 가 성립한다.

(증명). [그림 1.4]에서 $\overrightarrow{OE} \overset{\text{def}}{=} \vec{e}$가 단위벡터이고 $\overrightarrow{OP} \overset{\text{def}}{=} \vec{p} \neq \vec{0}$일 때 $c = \vec{p}\cdot\vec{e}$로 놓아 보자.

그러면

$$(\vec{p}-c\vec{e})\cdot\vec{e} = \vec{p}\cdot\vec{e} - c\vec{e}\cdot\vec{e} = c - c = 0$$

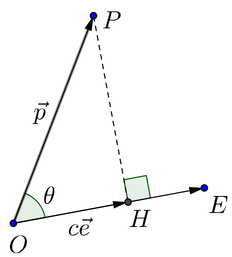


그림 1.4: 벡터의 사영

가 된다. 따라서 $\vec{p} - c\vec{e} \stackrel{\text{def}}{=} \overrightarrow{HP}$는 $\vec{e}$와 수직이며 $\vec{p}$와
$\vec{e}$ 사이의 각을 θ라고 할 때,
$\theta \le \pi/2$ 경우

$$\|\vec{p}\| \cos\theta = \overline{OH} = \|c\vec{e}\| = c \ge 0$$

가 될 것이다. 그러므로

$$\cos\theta = \frac{c}{\|\vec{p}\|} = \frac{\vec{p}\cdot\vec{e}}{\|\vec{p}\|} \tag{1.8}$$

가 된다. $\theta > \pi/2$일 때는 $\cos\theta < 0$, $c < 0$ 이므로 (1.8)은 역시 성립한다.

일반적으로, $\vec{p}$와 unit vector 아닌 nonzero 벡터 $\vec{q}$ 사이의 각 θ는 다음의 식을 만족할 것이다. ($\vec{e} = \frac{\vec{q}}{\|\vec{q}\|}$로 두면 된다.)

$$\cos\theta = \frac{\vec{p}\cdot\vec{q}}{\|\vec{p}\|\|\vec{q}\|}. \tag{1.9}$$

따라서 아래의 등식이 성립한다.

$$\vec{p}\cdot\vec{q} = \|\vec{p}\|\|\vec{q}\|\cos\theta. \tag{1.10}$$

(통상 (1.9)와 (1.10)은 코싸인의 가법정리를 써서 증명한다.) □

<u>단평</u> 1.10 물리학 책에서는 벡터의 내적을 (1.10)으로 정의하고, 이 정의로부터 $(a_1, a_2)\cdot(b_1, b_2) = a_1b_1 + a_2b_2$가 성립함을 코싸인의 가법정리를 써서 증명하는 것이 보통이다. ⊣

<u>단평</u> 1.11 벡터의 내적을 써서 코싸인 제2법칙을 쉽게 증명할 수 있다. 내적의 정의에 의하여 $\|\vec{p}-\vec{q}\|^2 = \|\vec{p}\|^2 + \|\vec{q}\|^2 - 2\,\vec{p}\cdot\vec{q}$가 성립함을 알 수 있는데, 이 식의 우변에서 $2\,\vec{p}\cdot\vec{q}$를 $2\,\|\vec{p}\|\,\|\vec{q}\|\cos\theta$로 대체하면 이것이 바로 코싸인 제2법칙이다. 이는 제4쪽에 있는 [그림 1.2]를 보면 알 수 있을 것이다. ⊣

<u>정의</u> 1.12 벡터 $\vec{p}$와 벡터 $\vec{q} \ne \vec{0}$의 출발점을 일치시키고 $\vec{p}$를 $\vec{q}$에($\vec{q}$를 품는 직선에) 사영(*project*)했을 때 얻어지는 벡터를

$$\mathrm{Proj}_{\vec{q}}\,\vec{p}$$

로 나타내고 $\vec{p}$의 $\vec{q}$에 대한 사영(*projection*)이라고 부른다. 또는 $\vec{p}$의 $\vec{q}$ 방향 성분(*component*)이라고 하고 $\vec{p} - \mathrm{Proj}_{\vec{q}}\,\vec{p}$를 p의 $\vec{q}$ 수직방향 성분이라고 한다. ⊣

<u>정리</u> *1.13* $Proj_{\vec{q}}\,\vec{p}$는 $\vec{q}$와 같은 방향, 혹은 정반대 방향의 벡터이며 $\vec{q}$를 품는 직선 상의 벡터 중에서는 $\vec{p}$와 가장 가깝다.† $Proj_{\vec{q}}\,\vec{p}$는 다음의 식으로 얻을 수 있다.

$$Proj_{\vec{q}}\,\vec{p} = \frac{\vec{q}\cdot\vec{p}}{\vec{q}\cdot\vec{q}}\,\vec{q} \tag{1.11}$$

(증명). [그림 1.4]를 보면 $\mathrm{Proj}_{\vec{q}}\,\vec{p} = \|\vec{p}\|\cos\theta\frac{\vec{q}}{\|\vec{q}\|}$ 이므로 여기에 (1.9)를 적용하여 (1.11)을 얻을 수 있다.

$\cos\theta$를 사용하지 않고도 다음과 같이 증명할 수 있다: $\mathrm{Proj}_{\vec{q}}\,\vec{p} = c\,\vec{q}$로 두면 $(\vec{p} - c\,\vec{q})\cdot\vec{q} = 0$가 성립한다. 이 방정식을 c에 대해서 풀면 $c = \frac{\vec{q}\cdot\vec{p}}{\vec{q}\cdot\vec{q}}$를 얻는다.

$\vec{q}$를 품는 직선 상의 벡터, 즉 $t\vec{q}$ 형태의 벡터 중에서는 $\mathrm{Proj}_{\vec{q}}\,\vec{p}$가 가장 $\vec{p}$에 가깝다는 것은 곧

$$f(t) = (t\vec{q} - \vec{p})\cdot(t\vec{q} - \vec{p}) = \|\vec{q}\|^2\,t^2 - 2(\vec{q}\cdot\vec{p})t + \|\vec{p}\|^2$$

로 두었을 때 f의 최솟값이 $t = \frac{\vec{q}\cdot\vec{p}}{\vec{q}\cdot\vec{q}}$에서 얻어진다는 것을 의미한다. 이 사실은 $f(t)$를 완전제곱으로 묶어 보면 곧 알 수 있다. □

<u>연습문제</u> 1.14 $\vec{p} = (3,1)$ 일 때 다음의 각 경우에 대하여 $\mathrm{Proj}_{\vec{q}}\,\vec{p}$를 계산하여라.

① $\vec{q} = (1,0)$, ② $\vec{q} = (-1,0)$,
③ $\vec{q} = (0,1)$, ④ $\vec{q} = (0,-1)$

또한 $\vec{p} = (-2,1)$ 일 때 위와 같은 요령으로 답하여라.

① $\vec{q} = (-1,2)$, ② $\vec{q} = (2,1)$ ⊣

1.3 평면 유클리드 기하학

유클리드 기하학의 많은 부분을 벡터로 작업할 수 있다. 이 섹션에서는 우선 $\mathbb{R}^2$ 에서의 도형만 다루기로 한다. 아래의 원칙들은 대단히 중요하므로 철저히 이해하여 어떤 상황에서든지 자유자재로 사용할 수 있어야 할 것이다. 아래에서 $\vec{u} \neq \vec{0} \neq \vec{v}$를 가정한다.

(*a*) $\vec{u}\cdot\vec{v} = 0$는 $\vec{u}$와 $\vec{v}$가 수직일 필요충분조건이다.

(*b*) 어떤 $k \in \mathbb{R}$에 대해서 $\vec{u} = k\vec{v}$라는 것은 (*i*) $k > 0$이면 두 벡터의 방향이 같을 필요충분조건이고 (*ii*) $k < 0$이면 두 벡터의 방향이 반대일 필요충분조건이다.

(*c*) 영벡터가 아니고 방향이 같거나 반대가 아닌 두 벡터 $\vec{u}$, $\vec{v}$에 대해서, 평면 내의 임의의 벡터 $\vec{w}$는 $\vec{w} = a\vec{u} + b\vec{v}$로 나타낼 수 있다. 각 $\vec{w}$에 대해서 이러한 $a, b \in \mathbb{R}$은 유일하게 결정된다. 이 선형조합(*linear combination*)의 계수(*coefficient*) a, b 중 하나, 혹은 모두가 음수 또는 0일 수 있다.

†물론 벡터들의 출발점을 일치시켰을 때의 얘기다.

(d) 영벡터가 아니고 방향이 같거나 반대가 아닌 두 벡터 $\vec{u}, \vec{v}$와 스칼라 a_1, b_1, a_2, b_2에 대하여 $a_1\vec{u} + b_1\vec{v} = a_2\vec{u} + b_2\vec{v}$가 성립한다면 $a_1 = a_2$이고 또한 $b_1 = b_2$이어야만 한다. 이것의 역은 당연히 성립한다.

(e) $\vec{u}$와 $\vec{v}$의 끝점을 잇는 선분의 $\alpha : \beta$ 내분점을 종점으로 하는 벡터 $\vec{w}$는

$$\begin{aligned}\vec{w} &= \frac{\beta\vec{u} + \alpha\vec{v}}{\alpha + \beta} \\ &= (1-t)\vec{u} + t\vec{v}, \quad \text{where } t = \frac{\alpha}{\alpha + \beta}\end{aligned}$$

이다. 통상 $\alpha > 0$, $\beta > 0$이므로 $0 < t < 1$이 만족된다.

위의 5 원칙들 중에 (a)와 (b)는 이제 충분히 설명되었다고 본다. (c)와 (d)는 평면에서 도형을 그려 생각하면 당연히 성립할 것으로 보이며 엄격한 증명은 조금 후에 제공할 것이다.

(e)에 대해서 설명을 하겠다. [그림 1.5]에 나타내었듯이 평면 상에 주어진 2개의 점 U, V를 잇는 선분 $\overline{UV}$ 위의 임의의 점 W는 실수 $0 < t < 1$에 대해서 벡터

$$\vec{w} \stackrel{\text{def}}{=} (1-t)\vec{u} + t\vec{v} \quad (*)$$

의 끝점으로 나타낼 수 있다는 것을 확인해 보자.

원점 O는 임의로 잡고 $\vec{u} \stackrel{\text{def}}{=} \overrightarrow{OU}$, $\vec{v} \stackrel{\text{def}}{=} \overrightarrow{OV}$, $\vec{w} \stackrel{\text{def}}{=} \overrightarrow{OW}$로 두었다. 그러면

$$\begin{aligned}\vec{w} &= \vec{u} + \overrightarrow{UW} \\ &= \vec{u} + \frac{\alpha}{\alpha+\beta}\overrightarrow{UV} \\ &= \vec{u} + t(\vec{v} - \vec{u}) \\ &= (1-t)\vec{u} + t\vec{v}\end{aligned}$$

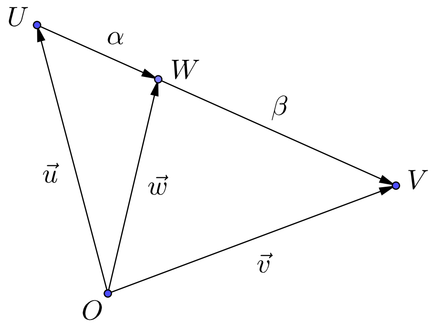


그림 1.5: 선분의 내분점

가 되어 $(*)$를 얻게 된다.

<u>정의</u> 1.15 $a_i, b_i, c_i \in \mathbb{R}$ $(i = 1, 2, 3)$일 때 아래의 왼쪽과 같은 2차원적 배열을 2×2-행렬(*matrix*)이라고 한다. (여기서 2×2는 "2 by 2"라고 읽는다.)

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \qquad \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

위의 오른쪽과 같은 2차원적 배열을 3×3-행렬이라고 한다.

위에서 정의한 2×2-행렬을 A라 했을 때 $a_1b_2-b_1a_2$를 이 A의 디터미넌트(*determinant*)라고 하고 $\det(A)$로 나타낸다. 위에 보인 3×3-행렬의 디터미넌트는 다음과 같이 계산하여 얻어지는 실수이다.

$$a_1 \det\begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} - b_1 \det\begin{pmatrix} a_2 & c_2 \\ a_3 & c_3 \end{pmatrix} + c_1 \det\begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$$

위 식의 가운데 항의 부호가 $-$인 것에 주목하라. ⊣

<u>보조정리</u> *1.16* 모든 $\vec{u} := (u_1, u_2) \neq \vec{0},\ v := (v_1, v_2) \in \mathbb{R}^2$에 대해서 다음의 두 명제는 서로 동등하다.

① $u_1v_2 - u_2v_1 = 0$

② $\vec{v} = k\vec{u}$가 성립하는 $k \in \mathbb{R}$가 존재한다.

(증명). 일반성의 손실없이 $u_1 \neq 0$를 가정하자.†

① ⇒ ②: $u_1v_2 - u_2v_1 = 0$이면 $v_2 = \frac{u_2}{u_1}v_1$이므로 $k = \frac{v_1}{u_1}$으로 두었을 때

$$k\vec{u} = \frac{v_1}{u_1}(u_1, u_2) = (v_1, v_2)$$

가 된다.

② ⇒ ①: $v_1 = ku_1,\ v_2 = ku_2$로 두면 $u_1v_2 - u_2v_1 = u_1(ku_2) - u_2(ku_1) = 0$를 얻는다. □

<u>연습문제</u> 1.17 [보조정리 1.16]은 $\vec{v} \neq \vec{0}$인 경우에는 어떻게 되겠는가? ⊣

<u>보조정리</u> *1.18* 다음의 두 명제들은 서로 동등하다.

① 미지수 x, y에 대한 다음의 *2*원*1*차 연립방정식이 유일한 해를 가진다.

$$u_1x + v_1y = w_1 \tag{1.12}$$

$$u_2x + v_2y = w_2 \tag{1.13}$$

② $u_1v_2 - u_2v_1 \neq 0$

(증명). ② ⇒ ①: ②가 성립하면 주어진 방정식에서 다음과 같이 y를 소거할 수 있다. $(1.12) \times v_2 - (1.13) \times v_1$를 계산하면

---

† $\vec{u} \neq \vec{0}$이면 $u_1 \neq 0$이거나 $u_2 \neq 0$일 것인데, 후자의 경우에는 전자의 경우에 대한 증명과 거의 같은 방법으로 증명할 수 있을 것이므로 전자의 경우에 대해서만 증명하여도 것이 충분할 것이다. 이러한 상황에서 '일반성의 손실없이'라는 표현을 흔히 사용하며 이를 영어로는 wlog(without loss of generality)라고 한다.

$$x = \frac{w_1v_2 - w_2v_1}{u_1v_2 - u_2v_1}$$

②로부터 $\vec{v} \neq \vec{0}$를 알 수 있으므로 wlog $v_1 \neq 0$를 가정하면 (1.12)로부터

$$\begin{aligned} y &= \frac{w_1 - u_1x}{v_1} = \frac{1}{v_1}\left(w_1 - u_1\frac{w_1v_2 - w_2v_1}{u_1v_2 - u_2v_1}\right) \\ &= \frac{1}{v_1} \cdot \frac{-w_1u_2v_1 + u_1w_2v_1}{u_1v_2 - u_2v_1} = \frac{-w_1u_2 + w_2u_1}{u_1v_2 - u_2v_1} \end{aligned}$$

의 해를 얻게 된다.

① ⇒ ②: $u_1v_2 - u_2v_1 = 0$를 가정하고 주어진 방정식의 해가 존재하지 않거나 또는 무한히 많이 존재함을 보이면 된다. 이것의 증명은 벡터공간과 행렬과 선형사상을 배운 뒤에는 깔끔하게 할 수 있으나 현재는 다음과 같이 여러 경우로 나누어 조금 지루하게 하는 수밖에 없다.

- $\vec{u} = \vec{0}$ 경우
  - $\vec{v} = \vec{0}$ 경우: $\vec{w} := (w_1, w_2) = \vec{0}$이면 모든 $(x, y) \in \mathbb{R}^2$가 방정식의 해가 되므로 해는 무한히 많다. $\vec{w} \neq \vec{0}$이면 (1.12)와 (1.13)의 좌변은 모두 0인데 우변은 적어도 하나가 0이 아니므로 해가 존재하지 않는다.
  - $\vec{v} \neq \vec{0}$ 경우: Wlog $v_1 \neq 0$라고 하자. 그러면 (1.12)로부터 $y = \frac{w_1}{v_1}$이 되어야 한다. 이제 (1.13)은 $v_2\frac{w_1}{v_1} = w_2$가 되는데 만일 이 등식이 성립한다면 모든 $x \in \mathbb{R}$에 대해서 $(x, \frac{w_1}{v_1})$가 방정식의 해가 되므로 해는 무한히 많다. 만일 $v_2\frac{w_1}{v_1} \neq w_2$라면 방정식의 해는 존재하지 않는다.
- $\vec{u} \neq \vec{0}$ 경우
  - $\vec{v} = \vec{0}$ 경우: 이 경우는 $\vec{u} = \vec{0}$이고 $\vec{v} \neq \vec{0}$인 경우와 똑같은 방법으로 방정식의 해가 무한히 많거나 존재하지 않음을 증명할 수 있다.
  - $\vec{v} \neq \vec{0}$ 경우: Wlog $v_1 \neq 0$를 가정한다. $u_1v_2 - u_2v_1 = 0$로부터 $v_2 = \frac{u_2}{u_1}v_1$가 얻어지므로 (1.13)은

    $$u_2x + \frac{u_2}{u_1}v_1y = w_2 \tag{$*$}$$

    가 된다.

    $u_2 = 0$인 경우에는 이 식의 좌변이 0이므로 우변 w_2가 0이어야 한다. 즉 $w_2 \neq 0$인 경우는 해가 존재하지 않고, $w_2 = 0$인 경우는 (1.13)은 항상 만족된다. 따라서 (1.12), 즉 $u_1x + v_1y = w_1$만 만족하면 해가 되는데 이러한 해는 무한히 많다—y는 임의의 실수로 두고 $x = w_1 - \frac{v_1y}{u_1}$으로 두면 된다.

마지막으로 $u_2 \neq 0 \neq u_1$ 인 경우가 남았다. 이 경우에는 $(*)$의 양변에 $\frac{u_1}{u_2}$를 곱하면 $u_1x+v_1y = \frac{u_1w_2}{u_2}$이 된다. 이 식은 (1.12)와 좌변은 동일하고 우변이 w_1 대신 $\frac{u_1w_2}{u_2}$로 바뀐 것이다. 따라서 $w_1 \neq \frac{u_1w_2}{u_2}$이면 해가 존재하지 않을 것이고 $w_1 = \frac{u_1w_2}{u_2}$이면 해가 무한히 많을 것이다. □

이제 [보조정리 1.16]과 [보조정리 1.18]을 사용하면 p8의 (c)가 성립함을 알 수 있다. p8의 (d)는 다음의 사실로부터 쉽게 보일 수 있다.

$$a_1\vec{u}+b_1\vec{v} = a_2\vec{u}+b_2\vec{v} \iff (a_1-a_2)\vec{u}+(b_1-b_2)\vec{b} = \vec{0},$$
$$a\vec{u}+b\vec{v} = \vec{0} \iff a=b=0$$

$u_1v_2-u_2v_1 \neq 0$는 두 0아닌 벡터 $\vec{u}, \vec{v}$가 갖추어야 할 중요한 성질인데 이를 선형독립(*linear independence*)이라고 하며 이는 앞으로 공부할 가장 중요하고 기본적인 개념이다.

<u>예제</u> 1.19 벡터를 이용하여 삼각형의 수심이 존재함을 증명해 보자.

$\triangle\ OAB$가 주어졌을 때 꼭짓점 A에서 변 OB로 수선 AC를 내리고, 꼭짓점 B에서 변 OA로 수선 BD를 내려 두 수선의 교점을 E라 한다. 선분 OE의 연장선이 변 AB와 만나는 점을 F라 하면 $OF \perp AB$이다.

이 사실을 벡터를 써서 증명하기 위하여 $\overrightarrow{OA} \overset{\text{def}}{=} \vec{a}$, $\overrightarrow{OB} \overset{\text{def}}{=} \vec{b}$로 놓고 이 두 벡터의 선형조합으로 $\overrightarrow{OE}$를 나타내 보자.

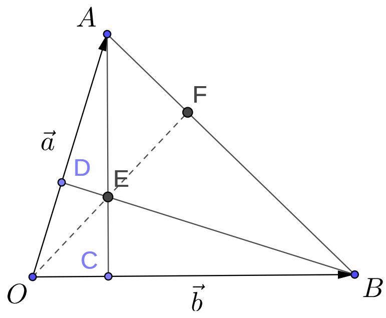


그림 1.6: 삼각형의 수심

$$\overrightarrow{OE} = t\vec{a}+s\vec{b}$$

가 되도록 $t, s \in \mathbb{R}$을 구하고 $\overrightarrow{OE}\cdot\overrightarrow{BA} = 0$를 보이면 된다.

$$(\overrightarrow{OE}-\vec{b})\cdot\vec{a} = 0,$$
$$(\overrightarrow{OE}-\vec{a})\cdot\vec{b} = 0$$

에서 $\overrightarrow{OE}$를 $t\vec{a}+s\vec{b}$로 바꾸면 t, s에 대한 2원1차 연립방정식이 아래와 같이 얻어진다.

$$((t\vec{a}+s\vec{b})-\vec{b})\cdot\vec{a} = 0 \quad\Rightarrow\quad t\,\|\vec{a}\|^2+s(\vec{a}\cdot\vec{b}) = \vec{a}\cdot\vec{b} \tag{1.14}$$
$$((t\vec{a}+s\vec{b})-\vec{a})\cdot\vec{b} = 0 \quad\Rightarrow\quad t(\vec{a}\cdot\vec{b})+s\|\vec{b}\|^2 = \vec{a}\cdot\vec{b} \tag{1.15}$$

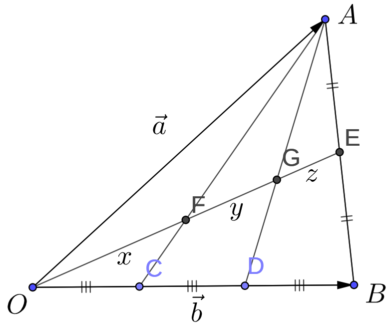
이 연립방정식의 해가 유일하게 존재함은 슈바르츠 부등식과 [보조정리 1.18]에 의하여 보장된다. 해를 구할 필요는 없고 아래의 등식이 성립함을 보이면 충분하다.

$$(t\vec{a}+s\vec{b})\cdot(\vec{b}-\vec{a}) = 0$$

위의 식은 (1.14)에서 (1.15)를 빼면 얻어진다. □

<u>예제</u> 1.20 [그림 1.7]에서 선분의 길이들의 비 $x : y : z$를 구해 보자.

그림 1.7: 예제 1.20



단, 점 C와 D는 변 OB의 3등분점이고 점 E는 AB의 중점이다. 또한 점 F와 G는 선분 OE와 선분 AC 및 선분 AD의 교점이다. 그리고 $x = \overline{OF}$, $y = \overline{FG}$, $z = \overline{GE}$이다.

이제

$$\overrightarrow{OF} = k_1\overrightarrow{OE}, \quad \overrightarrow{OG} = k_2\overrightarrow{OE} \quad (*)$$

가 되도록 실수 k_1, k_2를 구하면 $x : y : z = k_1 : k_2 - k_1 : 1 - k_2$로 계산된다. 우선 알 수 있는 것은 $\overrightarrow{OE} = \frac{\vec{a}+\vec{b}}{2}$라는 사실이다.

위의 $(*)$와 p9의 원칙 (e)를 이용하여 아래의 두 식을 얻는다.

$$k_1\frac{\vec{a}+\vec{b}}{2} = t_1\vec{a} + (1-t_1)\frac{\vec{b}}{3}, \quad \text{where } 0 < t_1 < 1,$$
$$k_2\frac{\vec{a}+\vec{b}}{2} = t_2\vec{a} + (1-t_2)\frac{2\vec{b}}{3}, \quad \text{where } 0 < t_2 < 1.$$

위의 두 식 모두의 좌변에서 $\vec{a}$와 $\vec{b}$의 계수가 같으므로 우변도 그러할 것이다. 따라서 $t_1 = \frac{1-t_1}{3}$으로부터 $t_1 = \frac{1}{4}$, $t_2 = \frac{2(1-t_2)}{3}$으로부터 $t_2 = \frac{2}{5}$를 얻는다.

그러므로 $k_1 = 2t_1 = \frac{1}{2}$, $k_2 = 2t_2 = \frac{4}{5}$이고 $x : y : z = \frac{1}{2} : \frac{4}{5} - \frac{1}{2} : 1 - \frac{4}{5} = 5 : 3 : 2$가 된다. ⊣

<u>연습문제</u> 1.21 [그림 1.7]에서 $\overline{AE} : \overline{EB} = 3 : 1$이고 C가 $\overline{OB}$의 중점이라고 했을 때 $\overrightarrow{OF} = \alpha\vec{a} + \beta\vec{b}$가 되는 $\alpha, \beta \in \mathbb{R}$을 구하여라. ⊣

<u>연습문제</u> 1.22 삼각형의 중심, 내심, 외심, 방심이 존재함을 벡터를 이용하여 각각 증명

하여라. ⊣

<u>연습문제</u> 1.23 두 직선이 수직일 필요충분조건은 기울기들의 곱이 -1 임을 증명하여라. (단, 수평선과 수직선의 경우는 예외이다.) ⊣

1.4 벡터의 외적

<u>정의</u> 1.24 $\mathbb{R}^3$ 내의 두 벡터 $\vec{p}$와 $\vec{q}$의 외적(*cross product*) $\vec{p}\times\vec{q}$는 다음과 같이 정의된다. $\vec{p}=(a_1,a_2,a_3)$, $\vec{q}=(b_1,b_2,b_3)$일 때

$$\vec{p}\times\vec{q}=(a_2b_3-a_3b_2,a_3b_1-a_1b_3,a_1b_2-a_2b_1). \tag{1.16}$$

위 식은 암기하기가 약간 까다로운데 행렬의 디터미넌트(*determinant*)를 이용하여 다음과 같이 간편하게 나타낼 수 있다. 먼저 $\mathbb{R}^3$의 기본벡터 3개를

$$\vec{\imath}\stackrel{\text{def}}{=}(1,0,0),\quad \vec{\jmath}\stackrel{\text{def}}{=}(0,1,0),\quad \vec{k}\stackrel{\text{def}}{=}(0,0,1)$$

로 정의한다. 그러면

$$(a_1,a_2,a_3)\times(b_1,b_2,b_3)=\det\begin{pmatrix}\vec{\imath}&\vec{\jmath}&\vec{k}\\ a_1&a_2&a_3\\ b_1&b_2&b_3\end{pmatrix}$$

가 되어 쉽게 암기할 수 있다. ⊣

<u>보조정리</u> *1.25 Cross product*의 주요 성질은 아래와 같다. 단 여기서 $\vec{p},\vec{q},\vec{r}\in\mathbb{R}^3$이고 $a\in\mathbb{R}$이다.

(*a*) $\vec{\imath}\times\vec{\jmath}=\vec{k}$, $\vec{\jmath}\times\vec{k}=\vec{\imath}$, $\vec{k}\times\vec{\imath}=\vec{\jmath}$.

(*b*) $\vec{p}\times\vec{p}=\vec{0}$.

(*c*) $\vec{p}\times\vec{q}=-\vec{q}\times\vec{p}$.

(*d*) $(\vec{p}\times\vec{q})\times\vec{r}\neq\vec{p}\times(\vec{q}\times\vec{r})$, *in general.*

(*e*) $\vec{p}\times(\vec{q}+\vec{r})=\vec{p}\times\vec{q}+\vec{p}\times\vec{r}$, *and* $(\vec{q}+\vec{r})\times\vec{p}=\vec{q}\times\vec{p}+\vec{r}\times\vec{p}$.

(*f*) $(a\vec{p})\times\vec{q}=a(\vec{p}\times\vec{q})=\vec{p}\times(a\vec{q})$.

(*g*) $(\vec{p}\times\vec{q})\times\vec{r}=(\vec{p}\cdot\vec{r})\vec{q}-(\vec{q}\cdot\vec{r})\vec{p}$.

(*h*) $\vec{p}\times\vec{q}$ *is perpendicular to both* $\vec{p}$ *and* $\vec{q}$.

(*i*) $\|\vec{p}\times\vec{q}\| = \|\vec{p}\|\,\|\vec{q}\|\,|\sin\theta|$.

(*j*) $|\vec{p}\cdot(\vec{q}\times\vec{r})| = |\vec{q}\cdot(\vec{r}\times\vec{p})| = |\vec{r}\cdot(\vec{p}\times\vec{q})|$ ⊣

<u>단평</u> 1.26 (c)와 (d)는 외적의 교환법칙과 결합법칙이 성립하지 않음을 말해 주고 있다. (g)의 좌변은 벡터3중곱(*vector triple product*)이라고 한다. 벡터의 3중곱에는 벡터3중곱과 (j)의 스칼라3중곱(*scalar triple product*)이 있으며, 그냥 3중곱이라고 하면 통상 후자를 뜻한다.

$\mathbb{R}^3$에서 3개의 벡터를 곱하여 스칼라를 얻으려면 먼저 2개의 벡터의 외적을 취하고 그 결과를 3번째 벡터와 내적시키는 수밖에 없는데, 가능한 6가지 조합의 결과는 항상 일정하다는 것이 (j)의 핵심이며, 더하여 중요한 사실은 이 값이 바로 이 3벡터들이 이루는 평행육면체의 부피라는 것이다.

(i)는 두 벡터의 외적으로 얻은 벡터의 길이는 두 벡터가 이루는 평행사변형의 면적이라는 것을 말하고 있다. 그리고 외적에서 가장 중요한 사실은 (h)로서 두 벡터의 외적은 원래의 두 벡터를 품는 평면에 수직하다는 것이다. 그리고 평면에 수직인 방향은 (서로 정반대인) 2개인데 벡터의 외적의 방향은 이 둘 중에 오른손법칙에 의하여 얻어지는 방향이 된다.

이상의 사실들은 모두 외적의 정의로부터 어렵지 않게 증명된다. ⊣

<u>연습문제</u> 1.27 $\mathbb{R}^3$의 3 벡터의 스칼라3중곱은 세 벡터를 행벡터로(혹은 열벡터로) 하는 정방행렬의 디터미넌트임을 보이시오. ⊣

<u>단평</u> 1.28 외적을 계산할 때 식 (1.16)을 꼭 암기하고 있어야 할 필요는 없다. 이 식을 사용하는 것보다는 보조정리 1.25의 $(a), (b), (c), (d), (f)$를 반복적용 하는 것이 더 쉽고 실수가 없을 것이다. 예를 들어

$$\begin{aligned}(1,2,-3)\times(-2,1,4) &= (\vec{\imath}+2\vec{\jmath}-3\vec{k})\times(-2\vec{\imath}+\vec{\jmath}+4\vec{k})\\ &= \big(\vec{0}+\vec{k}+4(-\vec{\jmath})\big)+\big(-4(-\vec{k})+\vec{0}+8\vec{\imath}\big)+\big(6\vec{\jmath}-3(-\vec{\imath})-\vec{0}\big)\\ &= 11\vec{\imath}+2\vec{\jmath}+5\vec{k}\end{aligned}$$

와 같이 계산할 수 있다. ⊣

<u>연습문제</u> 1.29 보조정리 1.25의 (i)를 증명하고, 아래의 $\vec{u}$와 $\vec{v}$를 두 변으로 하는 평행사변형의 면적을 계산하시오.

(1) $\vec{u} = (-3,6,1)$, $\vec{v} = (-1,-2,1)$

(2) $\vec{u} = (6,0,-1)$, $\vec{v} = (0,1,2)$

(3) $\vec{u} = (0, 0, 2)$, $\vec{v} = (8, -1, 0)$ ⊣

<u>연습문제</u> 1.30 보조정리 1.25의 (j)를 증명하고, 4개의 점 P, U, V, W가 아래와 같이 주어졌을 때 $\overrightarrow{PU}$, $\overrightarrow{PV}$, $\overrightarrow{PW}$를 (한 꼭짓점에서 모이는) 세 모서리로 하는 평행육면체의 부피를 계산하시오.

(1) $P = (1, 1, 1)$, $U = (-4, 2, 7)$, $V = (3, 5, 7)$, $W = (0, 1, 6)$

(2) $P = (1, 6, 1)$, $U = (-2, 4, 2)$, $V = (3, 0, 0)$, $W = (2, 2, -4)$ ⊣

1.5 직선과 평면의 벡터 방정식

지금까지 공부한 벡터의 기하적 의미를 이용하면 직선이나 평면의 방정식을 용이하게 만들 수 있다.

$\mathbb{R}^n$에서 벡터 $\vec{r}_0$의 끝점을 지나며 벡터 $\vec{v}$에 평행인 직선의 방정식은, 매개변수(*parameter*) t를 사용하여

$$\vec{r}(t) = \vec{r}_0 + t\vec{v}, \quad -\infty < t < \infty \tag{1.17}$$

로 나타낼 수 있다. 이 직선을 점들의 집합으로 보면

$$\{\vec{r}_0 + t\vec{v} \mid t \in \mathbb{R}\}$$

이다. (1.17)을 $\mathbb{R}^3$의 경우에 대해서 분석해 보자.

$$\begin{aligned} \vec{r} &= (x(t), y(t), z(t)) = (x, y, z), \\ \vec{r}_0 &= (x_0, y_0, z_0), \\ \vec{v} &= (v_x, v_y, v_z) \end{aligned}$$

로 두면 $(x, y, z) = (x_0 + tv_x, y_0 + tv_y, z_0 + tv_z)$이므로 약간의 계산을 거쳐

$$\frac{x - x_0}{v_x} = \frac{y - y_0}{v_y} = \frac{z - z_0}{v_z} \tag{1.18}$$

를 얻을 수 있다.

<u>연습문제</u> 1.31 $\mathbb{R}^2$에서 직선의 방정식 $y = ax + b$를 (1.17)의 형식으로 변환해 보시오. 역방향의 변환도 해 보시오. ⊣

$\mathbb{R}^2$에서 직선의 방정식을 나타내는 또 하나의 방법이 있다. 예를 들어 $\vec{r}_0 \stackrel{\text{def}}{=} (1, 2)$를 지나고 $\vec{u} \stackrel{\text{def}}{=} (-3, 2)$에 수직한 직선 상의 동점 (x, y)를 끝점으로 하는 벡터를 $\vec{r}$로

놓으면 $\vec{r} - \vec{r}_0$는 $\vec{u}$에 수직이므로 다음의 식들을 얻는다. 이들은 모두 동등하다.

$$(\vec{r} - \vec{r}_0) \cdot \vec{u} = 0,$$
$$((x, y) - (1, 2)) \cdot (-3, 2) = 0,$$
$$-3(x - 1) + 2(y - 2) = 0,$$
$$-3x + 2y = 1,$$
$$y = \tfrac{3}{2}x + \tfrac{1}{2}.$$

<u>연습문제</u> 1.32 직선 $y = \frac{1}{2}x - 1$을 $\vec{r} = r_0 + t\vec{v}$ 형식으로 변환하여 나타내어라. 또한 이 직선을 $(\vec{r} - \vec{r}_0) \cdot \vec{n} = 0$ 형식으로 변환하여 나타내어라. ⊣

$\mathbb{R}^3$에서 벡터 $\vec{r}_0$의 끝점을 지나며 벡터 $\vec{u}$, $\vec{v}$에 평행한 평면의 방정식은, 매개변수 t, s를 사용하여

$$\vec{r} = \vec{r}_0 + t\vec{u} + s\vec{v}, \quad -\infty < t < \infty,\ -\infty < s < \infty \tag{1.19}$$

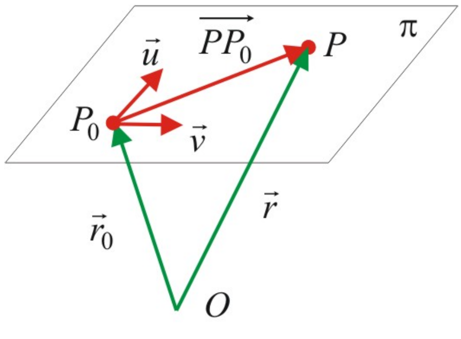
로 나타낼 수 있다. 이 평면을 점들의 집합으로 보면

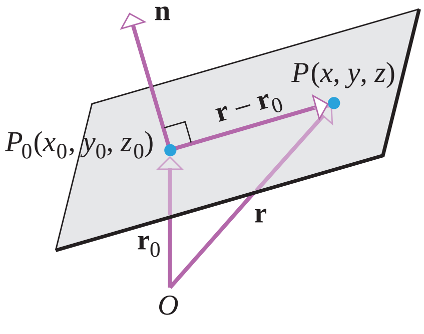
$$\{\vec{r}_0 + t\vec{u} + s\vec{v} \mid t, s \in \mathbb{R}\}$$

이다. $\mathbb{R}^3$에서 벡터 $\vec{r}_0$의 끝점을 지나며 벡터 $\vec{n}$에 수직인 평면의 방정식은

$$(\vec{r} - \vec{r}_0) \cdot \vec{n} = 0 \tag{1.20}$$

로 나타낼 수 있다. $\vec{n}$은 법선벡터(*normal vector*)라고 한다. 이 벡터 방정식 (1.20)은 $\mathbb{R}^3$에서는 평면을 나타내지만 $\mathbb{R}^2$에서는 앞서 보았듯이 직선을 나타낸다.





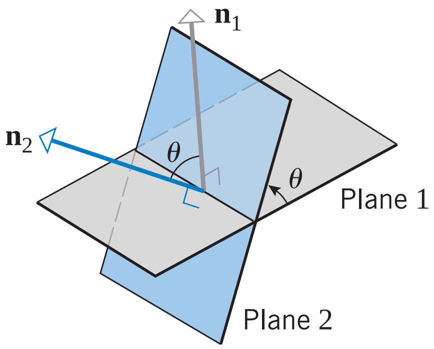
$\mathbb{R}^3$에서 평면을 나타내는 대표적인 4가지 방법으로 (1.19)와 (1.20), 그리고 $ax + by + cz = d$와 $z = ax + by + c$가 있다. 이 4가지 형태 간의 변환을 자유롭게 할 수 있어야 할 것이다.

- 벡터 (a,b,c)는 평면 $\pi : ax+by+cz=d$에 수직이다. 왜냐하면 π 위의 두 점 $P_1 : (x_1,y_1,z_1)$, $P_2 : (x_2,y_2,z_2)$를 잡았을 때 $\overrightarrow{P_1P_2}\cdot(a,b,c) = a(x_2-x_1)+b(y_2-y_1)+c(z_2-z_1) = (ax_2+by_2+cz_2)-(ax_1+by_1+cz_1) = d-d=0$이기 때문이다.
- 평면 $\vec{r}=\vec{r}_0+t\vec{u}+s\vec{v}$의 법선벡터는 $\vec{u}\times\vec{v}$로 구할 수 있다.
- 평면 $ax+by+cz=d$를 $\vec{r}=\vec{r}_0+t\vec{u}+s\vec{v}$ 형식으로 나타내려면 $\vec{u}$와 $\vec{v}$는 평면 상의 세 점 P_1, P_2, P_3를 마음대로 잡은 다음 $\vec{u}=\overrightarrow{P_1P_2}$, $\vec{v}=\overrightarrow{P_1P_3}$로 두면 된다. 또 하나의 방법은 법선벡터 (a,b,c)에 수직한 임의의 두 벡터를 잡는 것이다. 예를 들어 $\vec{u}=(b,-a,0)$, $\vec{v}=(0,c,-b)$로 두면 된다.

<u>연습문제</u> 1.33

(1) 다음 두 평면이 이루는 각도의 cosine을 구하여라. (힌트: 법선들이 이루는 각도)

$$2x-y+z=0,$$
$$x+2y-z=1.$$



(2) $Q=(1,1,1)$, $P=(1,-1,2)$, $N=(1,2,3)$이라고 하자. P를 지나고 $\overrightarrow{ON}$과 같은 방향인 직선과 Q를 지나고 $\overrightarrow{ON}$에 수직인 평면의 교점의 좌표를 구하여라. (힌트). 직선 $\vec{r}(t)=(x_0,y_0,z_0)+t(v_x,v_y,v_z)$와 평면 $ax+by+cz=d$의 교점은 t에 대한 방정식 $a(x_0+tv_x)+b(y_0+tv_y)+c(z_0+tv_z)=d$를 풀어서 구한다.

(3) 다음 세 점을 지나는 평면의 방정식을 (1) $\vec{r}=\vec{r}_0+t\vec{u}+s\vec{v}$ 형식으로 구하고, 또한 (2) $ax+by+cz+d=0$의 형식으로 구하여라.

① $(1,0,0)$, $(0,1,0)$, $(2,1,-1)$

② $(1,2,-1)$, $(-1,1,4)$, $(1,3,-2)$

(4) $\mathbb{R}^3$에서 주어진 점과 평면 사이의 최단거리를 구하는 방법을 설명하고, 점 $P :=(2,-1,4)$와 평면 $\pi := x-y+z=1$ 간의 거리를 구하여라. (힌트). P를 지나고 $\vec{n} := (1,-1,1)$ 방향인 직선과 π의 교점 Q를 구한 다음 선분 $\overline{PQ}$의 길이를 계산할 수도 있고, 다른 방법으로는 π 위의 임의의 점, 예를 들어 $P_0 := (1,0,0)$에서 P에 이르는 벡터 $\vec{u} := \overrightarrow{P_0P} = (1,-1,4)$를 $\vec{n}$에 사영하여 얻은 벡터 $\vec{v} := \mathrm{Proj}_{\vec{n}}\vec{u}$의 길이 $\|\vec{v}\| = |\vec{n}\cdot\vec{u}| \,/\, \|n\|$를 계산해도 된다. 후자를 이용하면 점 (x_1,y_1,z_1)과 평면 $ax+by+cz=d$ 간의 최단거리를 구하는 공식을 어렵지 않게 유도할 수 있다.

(5) $\mathbb{R}^2$와 $\mathbb{R}^3$의 각 경우에, 주어진 점과 직선 사이의 최단거리를 구하는 방법을 설명하여라. 그리고 직선 $\frac{x-1}{2} = y+2 = \frac{z}{3}$과 점 $\vec{p} := (1,2,1)$ 사이의 거리를 구하여라. (힌트). 직선의 방정식을 $\vec{r}(t) = \vec{r}_0 + t\vec{v}$ 형식으로 나타낸 다음 t에 대한 방정식 $(\vec{p}-\vec{r}(t))\cdot\vec{v} = 0$를 풀어 수선의 발의 좌표를 구한다. 다른 방법으로 직선 상의 점 $\vec{p}_0 := (1,-2,0)$을 잡고 $\vec{q} := \vec{p}-\vec{p}_0$를 $\vec{v}$에 사영한 벡터를 $\vec{q}$에서 빼어 얻은 벡터 $\vec{q} - \mathrm{Proj}_{\vec{v}}\vec{q}$의 길이를 구해도 된다. 수선의 발은 $\vec{p}_0 + \mathrm{Proj}_{\vec{v}}\vec{q}$로 구할 수 있다.

(6) 점 $P = (2,1,1)$에서 평면 $\vec{r} = (0,1,-1) + t(0,2,1) + s(-1,0,1)$에 내린 수선의 발의 좌표와 P에서 이 평면까지의 거리를 구하여라. (힌트). 평면의 방정식을 $ax+by+cz=d$ 형태로 얻는다. (a,b,c)는 $(0,2,1)\times(-1,0,1)$로 얻으면 되고, d는 $(x,y,z) = (0,1,-1)$을 대입하면 계산된다.

(7) ⓐ 두 평면 $\pi_1 \overset{\text{def}}{=} x-y+z=1$과 $\pi_2 \overset{\text{def}}{=} x-y+z=3$ 사이의 거리를 구하여라. (힌트). 점 $(1,0,0)\in\pi_1$과 π_2 간의 거리를 구한다. 혹은 각 평면에서 점 하나씩을 취하여 $P_1\in\pi_1$, $P_2\in\pi_2$로 두고 $\overrightarrow{P_1P_2}$를 법선벡터에 사영하여 얻은 벡터의 길이를 계산해도 된다.

ⓑ 두 평면 $\varphi_1 \overset{\text{def}}{=} ax+by+cz=d_1$과 $\varphi_2 \overset{\text{def}}{=} ax+by+cz=d_2$ 사이의 거리를 구하여라.

ⓒ π_1에 포함된 직선 2개 ℓ_1, ℓ_2를 구하여라. π_2에 포함된 직선 2개 m_1, m_2를 구하여라. ℓ_i와 m_j 간의 최단거리는 얼마인가? (힌트). 직선의 방향벡터는 평면의 법선벡터에 수직인 임의의 벡터를 잡으면 된다. 직선 상의 점은 평면 상의 임의의 점을 잡으면 된다.

ⓓ π_1 위의 점 P와 π_2 위의 점 Q가 주어졌을 때 벡터 $\overrightarrow{PQ}$를 이 평면들의 법선 벡터에 사영하여 얻은 벡터의 길이는 일정함을 증명하여라.

ⓔ $\mathbb{R}^3$에서 비꼬인 위치에 놓인 두 직선 ℓ_1과 ℓ_2 간의 최단거리를 구하는 방법을 간단히 설명하여라. (힌트). ℓ_i를 포함하는 평면 π_i를 잡되 이 두 평면이 평행하도록 한다. 즉 법선벡터를 공유하도록 한다. 법선벡터는 두 직선의

방향벡터들을 외적하여 얻는다.

(8) 점 P가 직선 $\ell_1 := (1+2t, -2+4t, 2-4t)$ 위를 움직이고 점 Q가 직선 $\ell_2 := (2+6t, 3+3t, 3+6t)$ 위를 움직일 때 선분 $\overline{PQ}$의 길이가 최소가 되는 점 P, Q의 좌표를 구하여라. (힌트). 두 직선의 방향 벡터를 각각 $\vec{v}_1 := (1, 2, -2)$, $\vec{v}_2 := (2, 1, 2)$로 두고, $\overrightarrow{PQ} = \big((2+2t) - (1+s), (3+t) - (-2+2s), (3+2t) - (2-2s)\big)$로 두었을 때 $\overrightarrow{PQ} \cdot \vec{v}_1 = 0$, $\overrightarrow{PQ} \cdot \vec{v}_2 = 0$가 성립해야 한다. ⊣

<u>연습문제</u> 1.34 $\mathbb{R}^3$에서 다음의 집합을 조건제시법을 써서 나타내 보아라.

(1) 두 벡터 $\vec{u}$와 $\vec{v}$를 인접한 두 변으로 가지는 평행사변형의 내부.

(2) 두 벡터 $\vec{u}$와 $\vec{v}$의 끝점을 잇는 선분. 그리고 $\vec{u}$와 $\vec{v}$를 두 변으로 하는 삼각형의 내부.

(3) 세 벡터 $\vec{u}$, $\vec{v}$, $\vec{w}$의 끝점들을 꼭짓점으로 가지는 삼각형의 내부.

(4) 4개의 벡터 $\vec{u}$, $\vec{v}$, $\vec{w}$, $\vec{p}$의 끝점들을 꼭짓점으로 하는 4면체의 내부. ⊣

2

벡터공간

2.1 부분공간, 기저, 차원

이 장은 상당히 추상적이므로 혹 완벽하게 이해하지 못하는 부분이 있을지라도 크게 염려할 필요는 없다고 본다.[†] 그러나 여기 나오는 대부분의 연습문제는 쉽게 풀 수 있을 정도의 직관력을 갖추는 수준까지는 공부해야 한다.

이 책에서 다루는 벡터공간은 주로 $\mathbb{R}^n$ 및 이것의 부분공간이 될 것이며, 가장 일반적인 경우에는 $\mathbb{R}$을 스칼라로 하는 내적공간이 될 것이다. 부분공간, 내적공간 등의 용어는 곧 설명하겠다.

정의 2.1 집합 $V \neq \varnothing$에 연산 $+$와 영벡터 $\vec{0}$와 스칼라곱 $\mathbb{R} \times V \to V$이 주어져 있고 이들이 [팩트 1.2]에 설명한 8개의 공리(또는 조건)들을 모두 만족할 때 순서쌍 $(V, \mathbb{R})$을 $\mathbb{R}$ 위의 벡터공간(*vector space*)이라고 한다.

V가 벡터공간일 때 $\cdot : V \times V \to \mathbb{R}$이 다음의 추가 조건을 만족하면

(1) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (교환법칙)

(2) $a(\vec{u} \cdot \vec{v}) = (a\vec{u}) \cdot \vec{v}$ (스칼라곱과 내적)

(3) $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$ (분배법칙)

(4) $\vec{u} \cdot \vec{u} \geq 0$ and $\vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$ (positive definiteness)

$\cdot$를 내적(*inner product*)이라고 하고 $(V, \cdot)$를 내적공간(*inner product space*)이라고 한다. 내적공간에서 두 벡터는 그들의 내적이 0일때 수직(*orthogonal, perpendicular*)하다고 정의한다. ⊣

내적의 공리, 즉 위에 보인 조건 (i)–(iv)들로부터 $a(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (a\vec{v})$를 증명할 수 있다.

[†]수학 비전공자의 가독성을 위하여 수학적 엄격성을 지키지 않은 정의나 표현들을 여기 저기 사용하였다.

$V = \mathbb{R}^n$ 일 때는 V 위의 내적으로 가장 널리 쓰이는 것은 $(u_1, \ldots, u_n)\cdot(v_1, \ldots, v_n) \stackrel{\text{def}}{=} \sum_i u_i v_i$ 로 정의되는 dot product이다. Dot product 이외의 다른 내적도 얼마든지 정의하여 유용하게 사용할 수 있지만[†] 이 책에서는 dot product 이외의 다른 내적은 공부하지 않는다.

<u>연습문제</u> 2.2 $\mathbb{R}^n$ 의 원소인 벡터들에 대해서는 ($k \neq 0$ and $\vec{u} \neq \vec{0}$)이면 $k\vec{u} \neq \vec{0}$가 성립함을 쉽게 확인할 수 있다. 일반적인 벡터 공간에서도 이 사실이 성립함을 증명하여라. (8개의 공리만 사용하여 증명해야 한다.) ⊣

—————— ∘ —————— ∘ ——————

여기서 잠시 벡터공간 $\mathbb{C}^n$ 의 내적에 대해서 알아보자. 이것은 스칼라가 $\mathbb{C}$ 인 벡터공간이지만 내적의 4번째 공리 positive definiteness는 여기서도 준수되어야 한다. Dot product를 $\mathbb{R}^n$ 에서와 같이 정의하면 (iv)가 만족되지 않는다. 예를 들어 $(1, 2i)\cdot(1, 2i) = 1 - 4 = -3 \not\geq 0$이다. $\mathbb{C}^n$ 에서의 dot product는, 내적이 갖추어야 할 조건 (iv)를 만족시키기 위하여 다음과 같이 정의된다.

$$(u_1, \ldots, u_n) \cdot (v_1, \ldots, v_n) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \overline{u}_i v_i \tag{2.1}$$

여기서 u_i 위의 overline은 켤레복소수(conjugate)를 의미한다. 이때 주의할 것은 이 정의는 내적의 4 공리 중에 (i)을 어기고 있다는 것이다. $\vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u}$가 아니고 $\vec{u}\cdot\vec{v} = \overline{\vec{v}\cdot\vec{u}}$이 성립한다. 또한 (ii) $a(\vec{u}\cdot\vec{v}) = \vec{u}\cdot(a\vec{v})$는 성립하지만 $a(\vec{u}\cdot\vec{v}) = (a\vec{u})\cdot\vec{v}$는 성립하지 않고, 대신에 $\bar{a}(\vec{u}\cdot\vec{v}) = (a\vec{u})\cdot\vec{v}$가 성립한다.

$\vec{q}$를 $\vec{p}$에 사영하여 얻은 벡터 $\text{Proj}_{\vec{q}}\,\vec{p}$는 (1.11)에서 $\frac{\vec{q}\cdot\vec{p}}{\vec{q}\cdot\vec{q}}\,\vec{q}$로 정의했었는데 많은 책에서는 이를 $\frac{\vec{p}\cdot\vec{q}}{\vec{q}\cdot\vec{q}}\,\vec{q}$로 정의한다. 거의 같아 보이지만 잘 보면 분자가 약간 다르다. $\mathbb{R}$ 위의 벡터공간이라면 내적에 교환법칙이 성립하니까 이렇게 써도 상관 없지만 $\mathbb{C}$ 위의 벡터공간에서는 후자를 쓰면 안 된다. 그 이유는, 후자의 정의를 사용한 경우에는

$$(\vec{p} - \text{Proj}_{\vec{q}}\vec{p}) \cdot \vec{q} = \vec{p}\cdot\vec{q} - \left(\frac{\vec{p}\cdot\vec{q}}{\vec{q}\cdot\vec{q}}\,\vec{q}\right)\cdot\vec{q} = \vec{p}\cdot\vec{q} - \frac{\overline{\vec{p}\cdot\vec{q}}}{\vec{q}\cdot\vec{q}}(\vec{q}\cdot\vec{q}) = \vec{p}\cdot\vec{q} - \overline{\vec{p}\cdot\vec{q}} \neq 0$$

이기 때문이다. (물론 이 값이 특수한 경우, 즉 $\vec{p}\cdot\vec{q} = 0$인 경우에는 0일 수 있다. 하지만 일반적으로는 0가 아니다.) $\text{Proj}_{\vec{q}}\,\vec{p}$의 정의를 원래대로 $\frac{\vec{q}\cdot\vec{p}}{\vec{q}\cdot\vec{q}}\,\vec{q}$를 사용하면 우리가 얻을 결과는 $\vec{p}\cdot\vec{q} - \overline{\vec{q}\cdot\vec{p}} = 0$가 되어 아무런 문제가 발생하지 않는다.

—————— ∘ —————— ∘ ——————

[†] 예를 들어 $(u_1, u_2) * (v_1, v_2) \stackrel{\text{def}}{=} 2u_1v_1 + u_2v_2$로 정의하면 $*$는 내적의 공리를 만족한다.

우리가 이 책에서 다루는 벡터공간은 $\mathbb{R}^n$ 이 대표적인데 이것 외에 우리가 다루는 벡터공간의 대표적인 예는 아래에 정의하는 "부분공간"이다.

<u>정의</u> 2.3 벡터공간의 부분집합 $V \neq \varnothing$는 다음의 성질을 만족할 때 부분공간(*subspace*) 이라고 한다.

(1) $(\forall \vec{p}, \vec{q} \in V)(\vec{p} + \vec{q} \in V)$,

(2) $(\forall \vec{p} \in V)(\forall c \in \mathbb{R})(c\vec{p} \in V)$.

V 의 성질 (i)을 두고 "덧셈에 더하여 닫혀있다(*closed*)."고 말하고 성질 (ii)를 두고는 "스칼라곱에 대하여 닫혀있다."고 말한다. ⊣

부분공간은 벡터공간이다. 왜냐하면 벡터공간의 공리 8개를 모두 만족하기 때문이다.

V가 벡터공간이고 U가 V 의 부분공간, W가 U 의 부분공간이면 W는 V 의 부분공간이다. 이 사실은 정의로부터 곧 알 수 있다.

부분공간을 "벡터공간의 부분집합으로서 그 자신이 벡터공간인 것"으로 정의해도 된다. 이 정의는 위의 정의와 동등하다.

<u>예제</u> 2.4 $V := \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$는 $\mathbb{R}^3$ 의 부분공간이다. 모든 $a, b \in \mathbb{R}$에 대해서 $U := \{(x, y, z) \in V \mid ax + by = 0\}$는 V 의 부분공간이다. 예를 들어 직선 $y = x$는 V 의 부분공간이다. $\{(0, 0, 0\}$는 U 의 부분공간이다. ⊣

<u>연습문제</u> 2.5 모든 $a, b, c \in \mathbb{R}$에 대해서 $V_2 := \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz = 0\}$는 $\mathbb{R}^3$ 의 부분공간임을 보여라. V_2의 부분공간을 찾아 보라. $d \in \mathbb{R}$일 때 평면 $ax+by+cz = d$는 $\mathbb{R}^3$ 의 부분공간인가? ⊣

부분공간은 항상 원점 $\vec{0}$을 원소로 가지게 된다. 그리고 V 의 가장 작은 부분공간은 $\{\vec{0}\}$ 이다. V 의 가장 큰 부분공간은 V 자신이다.

만일 $\vec{p} \neq \vec{0}$가 V 의 원소이면 $\{c\vec{p} \mid c \in \mathbb{R}\} \stackrel{\text{def}}{=} U \subseteq V$가 되어야 한다. $\vec{p} \in \mathbb{R}^2$, 또는 $\vec{p} \in \mathbb{R}^3$ 인 경우 U는 원점을 통과하는 직선이다. U는 $\mathbb{R}^n$ 의 부분공간임을 확인해 보라.

만일 $\vec{p}_1 \neq \vec{0} \neq \vec{p}_2$가 V 의 원소이면 $\{c_1\vec{p}_1 + c_2\vec{p}_2 \mid c_1, c_2 \in \mathbb{R}\} \stackrel{\text{def}}{=} U \subseteq V$가 되어야 한다. $\vec{p}_1, \vec{p}_2 \in \mathbb{R}^3$ 인 경우 U는 원점을 통과하는 평면이 되어야 하는데 예외적인 경우가 있으니 그것은 $\vec{p}_1 = k\vec{p}_2$ 인 스칼라 $k \neq 0$가 존재할 때이다. 이때는 U는 원점을 통과하는 직선이 된다.

<u>연습문제</u> 2.6 다음의 명제들 각각에 대해서 참인지 거짓인지 답하고 설명하여라.

(1) U와 W가 각각 V 의 부분공간이면 $U \cap W$ 도 V 의 부분공간이다.

(2) U와 W가 각각 V 의 부분공간이면 $U \cup W$ 도 V 의 부분공간이다. ⊣

연습문제 2.7 $\mathbb{R}^2$ 의 부분집합으로서 벡터합에 대해서는 닫혀있지만 스칼라곱에 대해서는 닫혀있지 않은 예를 들어라. $\mathbb{R}^2$ 의 부분집합으로서 스칼라곱에 대해서는 닫혀있지만 벡터합에 대해서는 닫혀있지 않은 예를 들어라. ⊣

정의 2.8 V 는 $\mathbb{R}$ 위의 벡터공간이고 $\vec{p}_1, \ldots, \vec{p}_m \in V$, $c_1, \ldots, c_m \in \mathbb{R}$, $m \geq 1$ 일 때

$$c_1\vec{p}_1 + \cdots + c_m\vec{p}_m$$

를 $\vec{p}_1, \ldots, \vec{p}_m$ 의 선형조합(*linear combination*)이라 하고[†] 이때 각 c_i 를 $\vec{p}_i$ 의 계수(*coefficient*)라고 한다. 계수들 중에 0이 아닌 것이 하나라도 있는 선형조합을 0 아닌 선형조합(*nontrival linear combination*)이라고 한다. 모든 계수가 0인 선형조합을 0-선형조합(*trivial linear combination*)이라고 한다.

$\vec{p}_1, \ldots, \vec{p}_m$ 의 모든 선형조합들이 이루는 집합을 이들이 이루는 스팬(*span*)이라 하고

$$\mathrm{Span}(\vec{p}_1, \ldots, \vec{p}_m)$$

으로 나타낸다. $U \subseteq V$ 일 때

$$\mathrm{Span}(U) \overset{\text{def}}{=} \{c_1\vec{p}_1 + \cdots c_m\vec{p}_m \mid m \in \mathbb{N},\ 1 \leq i \leq m,\ c_i \in \mathbb{R},\ \vec{p}_i \in U\}$$

로 정의한다.[‡]

$W = \mathrm{Span}(U)$ 일 때 W 는 U 에 의해서, 또는 U 로 스팬된다고 말한다. U 는 W 를 스팬한다고 말해기도 한다. 그리고 U 는 W 의 생성집합(*generating set*)이라고 한다. ⊣

연습문제 2.9

(1) $\mathrm{Span}(U) = \mathbb{R}^2$ 인 $U \subseteq \mathbb{R}^2$ 를 3개 찾되 U 의 원소의 개수가 각각 2, 3, 무한대가 되도록 해 보라. 원소의 개수를 1이 되도록 할 수 있겠는가?

(2) $\mathrm{Span}(U) = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 3z = 0\}$ 인 $U \subseteq \mathbb{R}^2$ 를 2개 이상 찾아 보라. ⊣

연습문제 2.10 V 가 벡터공간이고 $U \subseteq V$ 이면 $\mathrm{Span}(U)$ 는 V 의 부분공간이 되며 또한 이것은 U 를 품는 V 의 부분공간 중에서 최소임을 증명하여라. ⊣

연습문제 2.11 V 가 벡터공간이고 $U, U_1, U_2 \subseteq V$ 라 하자. 다음을 증명하시오.

[†] 선형조합은 때로는 이러한 표현(expression)을 의미하고 때로는 이 표현이 나타내는 벡터를 의미한다.

[‡] 그러니까 예를 들어 $\mathrm{Span}(\vec{p}_1, \vec{p}_2) = \mathrm{Span}(\{\vec{p}_1, \vec{p}_2\})$ 이다.

(i) $\mathrm{Span}(U) \supseteq U$ (확장성)

(ii) $U_1 \subseteq U_2 \ \Rightarrow\ \mathrm{Span}(U_1) \subseteq \mathrm{Span}(U_2)$ (순서보존성)

(iii) $\mathrm{Span}(U) = \mathrm{Span}(\mathrm{Span}(U))$ (멱등성, 선형조합의 선형조합은 선형조합이다: 3선의 법칙)

(힌트). (i), (ii)는 쉽다. (iii)을 위해서는 $\mathrm{Span}(\mathrm{Span}(U))$의 원소들은 일반적으로 어떤 형태일지 생각해 본다. ⊣

<u>연습문제</u> 2.12 V가 벡터공간이고 $\vec{v}_1, \vec{v}_2 \in V$, $t, s \in \mathbb{R}$, $s \neq 0$일 때 다음을 증명하여라.

$$\begin{aligned}\mathrm{Span}(\vec{v}_1, \vec{v}_2) &= \mathrm{Span}(\vec{v}_2, \vec{v}_1)\\ &= \mathrm{Span}(\vec{v}_1, \vec{v}_1 + \vec{v}_2)\\ &= \mathrm{Span}(\vec{v}_1, s\vec{v}_2)\\ &= \mathrm{Span}(\vec{v}_1, \vec{v}_2 + t\vec{v}_1)\end{aligned}$$

⊣

<u>정의</u> 2.13 벡터 $\vec{p}_1, \ldots, \vec{p}_m, (m \geq 1)$은 다음의 조건을 만족할 때 이들을 선형독립(*linearly independent*)이라고 한다. 선형독립의 부정은 선형종속(*linearly dependent*)이다.

$$(\forall c_1, \ldots c_m \in \mathbb{R})\big(c_1\vec{p}_1 + \cdots c_m\vec{p}_m = \vec{0} \ \Rightarrow\ c_1 = \ldots = c_m = 0\big)$$

$\vec{p}_1, \ldots, \vec{p}_m$는 선형독립(혹은 선형종속)이라고 말하기도 하고 $\{\vec{p}_1, \ldots, \vec{p}_m\}$는 선형독립(혹은 선형종속)이라고 말하기도 한다. (즉, 주어가 여러 개의 벡터일 수도 있고 하나의 집합일 수도 있다.)

벡터들의 (유한 혹은 무한)집합 U는 선형종속인 유한 부분집합을 가질때면이 선형종속이라고 정의한다. 그렇지 않으면 선형독립이다. 따라서 U는 모든 유한 부분집합이 선형독립일 때 선형독립이다. 공집합은 선형독립이다. ⊣

<u>단평</u> 2.14 $A := \{\vec{p}_1, \ldots, \vec{p}_m\}$의 0아닌 선형조합으로서 $\vec{0}$인 것이 존재하면이 A는 선형종속이다. A가 선형독립이면이 A의 선형조합으로서 $\vec{0}$인 것은 항상 0-선형조합이다. ⊣

<u>연습문제</u> 2.15 $\{\vec{u}_1, \ldots, \vec{u}_n\}$에 대해서 다음의 사실들을 증명하시오.

(1) $\vec{u}_i = \vec{0}$인 $i = 1, \ldots, n$이 존재하면 $\vec{u}_1, \ldots, \vec{u}_n$은 반드시 선형종속이다.

(2) $n = 1$인 경우: $\vec{u}_1 \neq \vec{0}$인 것은 $\vec{u}_1$이 선형독립일 필요충분조건이다.

(3) $n = 2$인 경우: $\vec{u}_1 \neq \vec{0}, \vec{u}_2$가 선형종속이면 $\vec{u}_2 = k\vec{u}_1$인 스칼라 k가 존재하며, 이것의 역도 성립한다.

(4) 선형독립인 집합의 부분집합은 선형독립이다. 선형종속인 집합의 초집합은 선형종속이다.

(5) $\{\vec{u}_1, \ldots \vec{u}_n\}$이 선형독립이면 "모든 $c_1, \ldots, c_n, d_1, \ldots, d_n \in \mathbb{R}$에 대해서

$$\textstyle\sum_{i=1}^n c_i\vec{u}_i = \sum_{i=1}^n d_i\vec{u}_i \;\Rightarrow\; (\forall i = 1, \ldots, n)(c_i = d_i)$$

이다." 그리고 "이것"의 역도 성립한다.

(6) $\{\vec{u}_1, \ldots \vec{u}_n\}$이 선형종속이면 이들 중의 어떤 $\vec{u}_i$는 나머지 벡터들 $\{\ldots \vec{u}_{i-1}, \vec{u}_{i+1}, \ldots\}$ 들의 선형조합이며 이것의 역도 성립한다. ($n = 2$인 경우에 대해서 먼저 증명해 보라.)

(7) $\vec{0}$ 아닌 벡터들이 모두 서로 수직하다면 이들은 선형독립이다. ⊣

<u>정의</u> 2.16 벡터공간 $V \neq \{\vec{0}\}$에 대해서, $U \subseteq V$가 다음의 조건을 만족할 때 U를 V의 기저(*basis*)라고 한다.

(1) U는 선형독립이다.

(2) $\mathrm{Span}(U) = V$.

즉, 기저는 선형독립인 생성집합이다. 벡터공간 $\{\vec{0}\}$의 기저는 $\varnothing$인 것으로 정한다. ⊣

<u>정의</u> 2.17 $\vec{0} \in \mathbb{R}^n$에서 i번째 성분만 1로 바꾸어 얻은 $\mathbb{R}^n$의 원소를 $\vec{e}_i \stackrel{\text{def}}{=} (\ldots, 0, 1, 0, \ldots)$ 라 하자. $\vec{e}_i$들은 기본벡터(*fundamental vector*)라고 하며 $\{\vec{e}_1, \ldots, \vec{e}_n\}$를 $\mathbb{R}^n$의 표준기저(*standard basis*)라고 한다. ⊣

<u>연습문제</u> 2.18

(1) $\mathbb{R}^2$의 기저를 표준기저 아닌 것으로 2개 이상 찾아 보시오.

(2) $\{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 3z = 0\}$의 기저를 2개 이상 찾아 보시오.

(3) 위의 (1)에서 기저의 원소의 개수가 1이 되도록 할 수 있는가? 원소의 개수가 3이 되도록 할 수 있는가? (2)에서는 어떠한가? ⊣

$\mathbb{R}^n$은 n 개의 원소로 이루어진 기저를 가짐을 쉽게 확인할 수 있다. 이 기저의 원소의 개수는 반드시 n이 되어야 한다는 사실을 곧 증명할 것이다. 아래의 정리는 모든 벡터공간의 기저에 대한 중요한 정리이다. 일반적인 경우에 대한 증명은 이 책의 범위를 벗어나므로 생략한다.

<u>정리</u> *2.19* 모든 벡터공간 V는 기저를 가진다. 그리고 벡터공간의 기저의 원소의 개수는 일정하다. 즉 U와 W가 V의 기저이면 U의 원소의 개수는 W의 원소의 개수와 같다.

(증명). 증명은 이 책의 범위를 넘으므로 생략한다. 단, V가 유한개의 벡터로 스팬되는 경우에 대한 증명은 [정리 2.22]에 제시하였다. □

정의 2.20 벡터공간의 기저의 원소의 개수를 그 벡터공간의 차원(*dimension*)이라고 한다. 그리고 벡터공간 V의 차원은 $\dim(V)$로 나타낸다. $\{\vec{0}\}$의 차원은 0인 것으로 정의한다. ⊣

보조정리 *2.21* V는 n개의 원소로 이루어진 집합 $W \overset{\text{def}}{=} \{\vec{w}_1, \ldots, \vec{w}_n\}$로 스팬되는 벡터공간이고, $U \overset{\text{def}}{=} \{\vec{u}_1, \ldots, \vec{u}_n\} \subseteq V$는 선형독립이면 U와 W는 각각 V의 기저를 이룬다.

(증명). $\operatorname{Span}(U) = V$이면 U는 V의 기저이다. 우리는 모든 $i = 1, \ldots, n$에 대하여

$$\operatorname{Span}(\vec{u}_1, \ldots, \vec{u}_i, W_i) = V$$
$$\text{단, } W_i \subseteq W,\ |W_i| = n - i$$

가 성립함을 i에 대한 수학적 귀납법을 사용하여 보일 것이다.

$\vec{u}_1$을 다음과 같은 선형조합으로 나타낸다.

$$\vec{u}_1 = a_{11}\vec{w}_1 + \cdots + a_{1n}\vec{w}_n$$

$\vec{u}_1 \neq \vec{0}$이므로 a_{1i} 중에 적어도 하나는 0이 아니다. 일반성의 손실 없이 $a_{11} \neq 0$를 가정하면

$$\vec{w}_1 = \frac{1}{a_{11}}\vec{u}_1 + \left(-\frac{a_{12}}{a_{11}}\right)\vec{w}_2 + \cdots + \left(-\frac{a_{1n}}{a_{11}}\right)\vec{w}_n$$

와 같이 $\vec{w}_1$을 $\vec{u}_1, \vec{w}_2, \ldots, \vec{w}_n$의 선형조합으로 나타낼 수 있다. 따라서

$$\operatorname{Span}(\vec{u}_1, \vec{w}_2, \ldots \vec{w}_n) = V \tag{2.2}$$

가 된다. (보충설명 : $W' := \{\vec{u}_1, \vec{w}_2, \ldots \vec{w}_n\}$로 놓으면 $\vec{w}_1 \in \operatorname{Span}(W')$이고 따라서 $W \subseteq \operatorname{Span}(W')$이다. 여기서 [연습문제 2.11]을 적용하면

$$V = \operatorname{Span}(W) \subseteq \operatorname{Span}(\operatorname{Span}(W')) = \operatorname{Span}(W')$$

가 성립함을 알 수 있다.)

이제 귀납가설로

$$\operatorname{Span}(\vec{u}_1, \ldots, \vec{u}_{i-1}, \vec{w}_i, \ldots \vec{w}_n) = V \tag{2.3}$$

를 가정하자. 그러면 $\vec{u}_i$를 다음과 같은 선형조합으로 나타낼 수 있다.

$$\vec{u}_i = a_{i1}\vec{u}_1 + \cdots + a_{i,i-1}\vec{u}_{i-1} + a_{ii}\vec{w}_i + \cdots + a_{in}\vec{w}_n$$

U가 선형독립이므로 $a_{ii}, \ldots, a_{in}$ 중에 적어도 하나는 0이 아니다. 일반성의 손실 없이 $a_{ii} \neq 0$를 가정하면

$$\vec{w}_i = \left(-\frac{a_{i1}}{a_{ii}}\right)\vec{u}_1 + \cdots + \left(-\frac{a_{i,i-1}}{a_{ii}}\right)\vec{u}_{i-1} + \frac{1}{a_{ii}}\vec{u}_i + \left(-\frac{a_{i,i+1}}{a_{ii}}\right)\vec{w}_{i+1} + \cdots + \left(-\frac{a_{1n}}{a_{ii}}\right)\vec{w}_n$$

와 같이 $\vec{w}_i$를 $\vec{u}_1, \ldots, \vec{u}_i, \vec{w}_{i+1}, \ldots, \vec{w}_n$의 선형조합으로 나타낼 수 있다. 이 사실과 귀납가설 (2.3)에 의하여 $\vec{w}_1, \ldots, \vec{w}_{i-1}, \vec{w}_i$ 각각은 모두 $\vec{u}_1, \ldots, \vec{u}_i, \vec{w}_{i+1}, \ldots, \vec{w}_n$의 선형조합으로 나타낼 수 있으므로 결국

$$\mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_i, \vec{w}_{i+1}, \ldots \vec{w}_n) = V \tag{2.4}$$

가 성립하게 된다.

(2.2)와 (2.3) $\Rightarrow$ (2.4)를 보였으므로 수학적 귀납법에 의하여 $\mathrm{Span}(U) = V$이다. 따라서 U는 V의 기저이다.

이제 W가 V의 기저임을 보이기 위하여는 W가 선형독립임을 보이면 된다. 만일 W가 선형독립이 아니라면 W의 어떤 원소는 다른 원소들의 선형조합이다. 일반성의 손실 없이 이 원소를 w_n이라 하면 $\{\vec{w}_1, \ldots, \vec{w}_{n-1}\}$은 V를 스팬한다. 앞서 보인 결과에 의하여 $\{\vec{u}_1, \ldots, \vec{u}_{n-1}\}$은 V를 스팬한다. 이것은 U가 선형독립이라는 가정에 모순이다. □

<u>정리</u> *2.22* 유한개의 벡터로 스팬되는 벡터공간 V의 두 부분집합 $W \stackrel{\text{def}}{=} \{\vec{w}_1, \ldots, \vec{w}_n\}$과 $U \stackrel{\text{def}}{=} \{\vec{u}_1, \ldots, \vec{u}_m\}$이 모두 기저라면 $n = m$이다.

(증명). 일반성의 손실없이 $m > n$을 가정하고 모순을 유도하면 된다.

$U' := \{\vec{u}_1, \ldots, \vec{u}_n\}$이 선형독립이므로 [보조정리 2.21]에 의하여 U'은 V의 기저이다. 따라서 $\vec{u}_{n+1} \in \mathrm{Span}(U')$이다. 이것은 U의 선형독립성에 모순된다. □

<u>보조정리</u> *2.23* V는 벡터공간, $U \subseteq V$는 선형독립이고 $\vec{v} \notin Span(U)$이면 $\{\vec{v}\} \cup U$도 선형독립이다.

(증명). $\{\vec{v}\} \cup U$가 선형종속이라고 가정하고 모순을 유도하겠다. 가정으로부터 $\{\vec{v}\} \cup U$의 유한 부분집합 W의 0 아닌 선형조합으로서 $\vec{0}$를 이루는 것이 존재한다. U가 선형독립이므로 W는 $\vec{v}$를 원소로 가져야만 한다. 그리고 이 선형조합에서 $\vec{v}$의 계수는 0이 아니어야 한다. 즉,

$$a_1\vec{u}_1 + \cdots + a_k\vec{u}_k + b\vec{v} = \vec{0}, \quad b \neq 0$$

로 둘 수 있다. 그러면 $\vec{v} = \left(-\frac{a_1}{b}\right)\vec{u}_1 + \cdots + \left(-\frac{a_k}{b}\right)\vec{u}_k \in \mathrm{Span}(U)$ 라는 모순을 얻는다. ⊣

<u>정의</u> 2.24 V를 벡터공간이라 하자. 벡터들의 집합 $U \subseteq V$는 다음의 두 성질을 만족할 때 V 안에서 극대독립(*maximally independent*)이라고 한다.

(1) U는 선형독립.

(2) $U \subsetneq U' \subseteq V$인 모든 U'은 선형종속. ⊣

<u>연습문제</u> 2.25 $(1,0),(1,1)$은 $\mathbb{R}^2$ 안에서 극대독립임을 보여라. $V := \{(x,y,z) \in \mathbb{R}^3 \mid x - y + z = 0\}$로 두었을 때 V 안에서 극대독립인 집합을 2개 구해 보아라. ⊣

<u>정리</u> *2.26* V가 벡터공간일 때 $U \subseteq V$에 대한 아래의 *3* 조건들은 모두 서로 등등하다.

① U는 V의 기저이다.

② U는 V 안에서 극대독립이다.

③ V의 모든 원소는 U의 원소들의 선형조합으로 유일하게 나타낼 수 있다. ⊣

(증명). ① $\Rightarrow$ ② $\Rightarrow$ ③를 보이고 이어서 $\neg$① $\Rightarrow$ $\neg$③을 보인다. 이하 생략. □

<u>보조정리</u> *2.27* V가 유한개의 벡터로 스팬되는 벡터공간이고 $U \subseteq V$가 $Span(U) = V$를 만족하면 U의 어떤 부분집합은 V의 기저가 된다.

(증명). 일반성의 손실 없이 $V \neq \{\vec{0}\}$를 가정한다. $\vec{u}_1 \in U - \{\vec{0}\}$를 취한다.

만일 $\mathrm{Span}(\vec{u}_1) = V$이면 V는 $\{\vec{u}_1\} \subseteq U$을 기저로 가진다.

그렇지 않다면 $\mathrm{Span}(*)$의 순서보존성(연습문제 2.11)을 이용하여 $\mathrm{Span}(\vec{u}_1) \not\supseteq U$를 보일 수 있고, 따라서 $\vec{u}_2 \in U - \mathrm{Span}(\vec{u}_1)$을 취할 수 있다. $\{\vec{u}_1, \vec{u}_2\}$는 [보조정리 2.23]에 의하여 선형독립이다.

만일 $\mathrm{Span}(\vec{u}_1, \vec{u}_2) = V$이면 V는 $\{\vec{u}_1, \vec{u}_2\}$를 기저로 가진다. 그렇지 않다면 $\vec{u}_3 \in U - \mathrm{Span}(\vec{u}_1, \vec{u}_2)$를 취한다.

이런 식으로 선형독립인 $\vec{u}_1, \ldots, \vec{u}_i \in U$를 얻었을 때

(경우 1). $\mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_i) = V$이면 V는 $\{\vec{u}_1, \ldots, \vec{u}_i\}$를 기저로 가진다.

(경우 2). 아니면 $\vec{u}_{i+1} \in U - \mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_i)$를 취한다.

[보조정리 2.21]에 의하여, 이러한 과정 중에 $\mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_k) = V$인 $k \leq n$이 반드시 존재하며, 이때 $\{\vec{u}_1, \ldots, \vec{u}_k\} \subseteq U$는 V의 기저이다. □

<u>따름정리</u> *2.28* 벡터공간 V가 그것의 어떤 원소들 $\vec{u}_1, \ldots, \vec{u}_k$에 대하여 $V = Span(\vec{u}_1, \ldots, \vec{u}_k)$라면 V는 u_i들 중의 일부(또는 전부)로 구성된 기저를 가진다.

(증명). V는 유한 개의 벡터로 스팬되는 벡터공간이므로 [보조정리 2.27]을 적용할 수 있다. □

보조정리 *2.29* $V \subseteq \mathbb{R}^n$가 벡터공간이고 $U \subseteq V$가 선형독립이면 U의 어떤 초집합은 V의 기저가 된다.

(증명). [보조정리 2.27]의 증명을 흉내낸다. □

단평 2.30 [따름정리 2.28]은 큰 생성집합에서 원소를 빼내어 선형독립성을 얻는 것이고, [보조정리 2.29]는 작은 선형독립 집합에 원소를 더하여 생성성을 얻는 것이다. 각각 기저의 downward 구성, upward 구성이라고 불러도 될 것이다.

모든 벡터공간이 기저(Hamel basis)를 가진다는 정리의 증명은 [정리 2.26]에 의하여 극대독립 부분집합을 구성하는 것으로 환원된다. 이 구성에는 통상 조른의 보조정리를 사용하는데 이것은 본질적으로 서수의 클라스(class of all ordinals)에서의 귀납과 같다. 귀납에는 원소 하나를 추가하는 단계와, 지금까지 얻어진 것들의 합집합을 취하는 두 단계가 있는데 [보조정리 2.27]은 이 중에 원소 하나를 추가하는 단계에 대응된다고 보면 될 것이다. ⊣

보조정리 *2.31* W가 벡터공간 $V \subseteq \mathbb{R}^n$의 부분공간이고 이 둘의 차원이 같다면, 즉 $W \subseteq V$이고 $\dim(W) = \dim(V)$이면 $W = V$이다.† 따라서 m-차원 벡터공간의 선형독립 부분집합 U의 원소의 개수가 m이면 U는 이 벡터공간의 기저이다.

(증명). [보조정리 2.21]을 이용한다. □

연습문제 2.32 세 벡터 $\vec{a}, \vec{b}, \vec{c}$가 선형종속이지만 이들 중 임의의 두 벡터가 선형독립인 예를 들라. ⊣

연습문제 2.33 이 문제에서 등장하는 모든 벡터는 $\vec{0}$가 아니다.

(*i*) $\mathbb{R}^2$에서 $\vec{a} \cdot \vec{c} = 0$, $\vec{b} \cdot \vec{c} = 0$이면 $\vec{a}$는 $\vec{b}$의 스칼라곱이 되어야 함을 보여라.

(*ii*) $\mathbb{R}^3$에 선형독립인 두 벡터 $\vec{c}_1, \vec{c}_2$가 있고 $\vec{a} \cdot \vec{c}_i = \vec{b} \cdot \vec{c}_i = 0$, $(i = 1, 2)$일 때도 같은 결론을 얻어 보아라.

(힌트). (i)의 결론은 $\mathbb{R}^2$가 아닌 임의의 2차원 내적공간에 대해서도 증명할 수 있다. $\vec{b}, \vec{c}$는 기저를 이루므로 $\vec{a} = \alpha\vec{b} + \beta\vec{c}$로 놓고 양변에 $\vec{c}$를 내적하여 $\beta = 0$임을 보인다. (ii)는 먼저 $\vec{c}_1$, $\vec{c}_2$, $\vec{c}_1 \times \vec{c}_2$가 $\mathbb{R}^3$의 기저를 이룬다는 것을 증명하고(기하적 직관 대신

---

† 일반적인 경우에는 무한차원 벡터공간이 존재하며 이때는 전체공간과 차원이 같은 부분공간이 전체공간의 진부분집합이 될 수도 있음을 말해 둔다.

$\vec{c}_1 \times \vec{c}_2$ 를 $\vec{c}_1$ 과 $\vec{c}_2$ 의 선형조합으로 두었을 때 [팩트 2.34]와 슈바르츠 부등식을 사용하여 모순 유도), 이를 이용하여 $\vec{a}$와 $\vec{b}$를 각각 $\vec{c}_1, \vec{c}_2, \vec{c}_1 \times \vec{c}_2$ 의 선형조합으로 나타내었을 때 $\vec{c}_1, \vec{c}_2$ 의 계수들이 모두 0임을 보임으로써 증명이 가능하다.

외적을 정의할 수 없는 일반적인 3차원 공간에 대해서는 이를 증명할 수는 있지만 현재까지 공부하여 얻은 지식으로는 어렵다. 이 문제는 다음 절에서 공부할 직교기저 (orthogonal basis)의 개념을 쓰면 해결된다. ⊣

팩트 2.34 x, y에 대한 연립방정식

$$ax + by = e,$$
$$cx + dy = f$$

는 [보조정리 1.18]에 의하여 $ad - bc \neq 0$ 이면이 유일한 해를 가진다. 이때의 해는

$$x = \frac{\det\begin{pmatrix} e & b \\ f & d \end{pmatrix}}{\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}}, \qquad y = \frac{\det\begin{pmatrix} a & e \\ c & f \end{pmatrix}}{\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}}$$

임을 단순한 계산에 의하여 확인할 수 있다. 따라서 $e = f = 0$ 인 경우에는 $(x, y) = (0, 0)$ 가 해이다.

$ad - bc = 0$ 이면 해가 존재하지 않거나 무한히 많이 존재한다. ⊣

보조정리 *2.35* $\vec{v}, \vec{w}$ 가 벡터이고 a_1, a_2, b_1, b_2 가 스칼라일 때, $\vec{p} \overset{\text{def}}{=} a_1\vec{v} + b_1\vec{w}$ 와 $\vec{q} \overset{\text{def}}{=} a_2\vec{v} + b_2\vec{w}$ 는

(*i*) $a_1b_2 - b_1a_2 = 0$ 이면 선형종속이다.

(*ii*) $a_1b_2 - b_1a_2 \neq 0$ 이고 $\vec{v}$, $\vec{w}$ 가 선형독립이면 선형독립이다.

(증명). (i) : $b_2\vec{p} - b_1\vec{q} = \vec{0}$.

(ii) : $x\vec{p} + y\vec{q} = \vec{0}$ 로 놓고 $x = y = 0$ 임을 보이면 된다. □

팩트 2.36 V 가 차원 n 의 벡터공간일 때 다음 사실들이 성립한다.

(*i*) n 개의 선형독립인 벡터들은 V 의 기저를 이룬다.

(*ii*) $m > n$ 개의 벡터들은 항상 선형종속이다.

(*iii*) $U \subseteq W$ 가 V 의 부분공간이면 $U \subsetneq W \;\Leftrightarrow\; \dim(U) < \dim(W)$ 이다. ⊣

2.2 사영과 직교기저

[정의 1.12]는 사영(projection)을 기하적으로 정의한 것이다. 대수적으로 정의한다면 식 $\mathrm{Proj}_{\vec{q}}\,\vec{p} = \frac{\vec{q}\cdot\vec{p}}{\vec{q}\cdot\vec{q}}\,\vec{q}$을 사용하면 되겠다. 이것은 1차원 부분공간 $\mathrm{Span}(\vec{q})$에 대한 사영이다. 이제 이 정의를 확장하여 일반적인 부분공간 $U \subseteq \mathbb{R}^n$에 대한 사영 $\mathrm{Proj}_U\vec{p}$를 정의해 보자. 이 정의를 이해를 돕기 위하여 다음의 문제를 풀어 보기 바란다.

<u>연습문제</u> 2.37 벡터 공간의 임의의 두 원소 $\vec{p}$, $\vec{w}$에 대해서

$$\big(\vec{p} - \mathrm{Proj}_{\vec{w}}\vec{p}\big) \ \perp \ \vec{w}$$

가 성립함을 보여라. ⊣

<u>정의</u> 2.38 $U \subseteq V$가 부분공간일 때 $\vec{p} \in V$의 U에 대한 사영(*projection*)은 다음의 조건을 만족하는 벡터 $\vec{u}$이다.

$$\vec{u} \in U \text{ and } \Big((\vec{p} - \vec{u}) \cdot \vec{w} = 0 \text{ for all } \vec{w} \in U\Big) \tag{2.5}$$

이러한 벡터 $\vec{u}$를 $\mathrm{Proj}_U\vec{p}$로 나타낸다.

$W \subseteq V$의 U에 대한 사영은 다음과 같이 정의된다: $\mathrm{Proj}_U(W) \stackrel{\text{def}}{=} \{\mathrm{Proj}_U(\vec{p}) \mid \vec{p} \in W\}$. ⊣

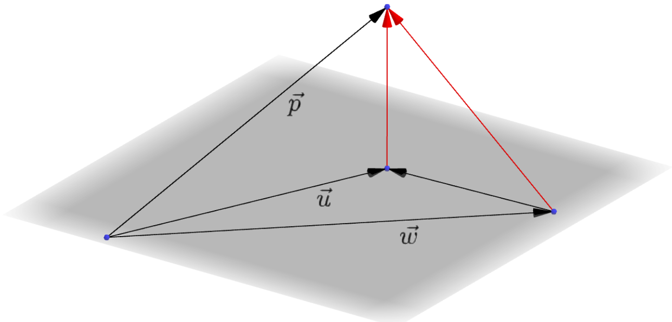
<u>예제</u> 2.39 $\vec{p} := (1,2,3) \in \mathbb{R}^3$를 부분공간 $U = \{(x,y,0) \mid x,y \in \mathbb{R}\} \subseteq \mathbb{R}^3$에 사영하면 $\vec{u} := (1,2,0)$을 얻는다. 이때 $\vec{p} - \vec{u} = (0,0,3)$이 되어 모든 $(x,y,0) \in U$에 수직이다. ⊣

이제 우리가 처음으로 해야 할 일은 $\mathrm{Proj}_U\vec{p}$가 유일하게 존재한다는 사실을 증명하는 것이다. 유일성을 보이기는 쉽다. $\vec{u}_1$과 $\vec{u}_2$가 (2.5)에서 $\vec{u}$의 역할은 한다면 모든 $\vec{w} \in U$에 대하여, $(\vec{p} - \vec{u}_1) \cdot \vec{w} = 0$ and $(\vec{p} - \vec{u}_2) \cdot \vec{w} = 0$이고, 따라서 $(\vec{u}_1 - \vec{u}_2) \cdot \vec{w} = 0$가 된다. 이제 $\vec{w} = \vec{u}_1 - \vec{u}_2$로 놓으면 $(\vec{u}_1 - \vec{u}_2) \cdot (\vec{u}_1 - \vec{u}_2) = 0$이므로 $\vec{u}_1 = \vec{u}_2$를 얻을 수 있다. 존재성의 증명은 썩 쉽지 않으며 수직 기저의 개념을 이용하면 비로소 증명된다. 유일·존재성의 증명에 앞서 (이를 가정하고) 다음의 보조정리를 공부한다.

<u>보조정리</u> *2.40* $U \subseteq \mathbb{R}^n$이 벡터공간이고 $\vec{p} \in \mathbb{R}^n$일 때 $\vec{u} \stackrel{\text{def}}{=} \mathit{Proj}_U\vec{p}$는 다음의 조건을 만족한다.

$$(\forall \vec{w} \in U)\Big(\|\vec{p} - \vec{w}\| \geq \|\vec{p} - \vec{u}\|\Big) \tag{2.6}$$

또한, 역으로 (2.6)을 만족하는 $\vec{u} \in U$는 (2.5)를 만족한다. 그러므로 $\mathit{Proj}_U\vec{p}$는 U의 원소 중에서 $\vec{p}$와 가장 근사(*approximate*)한 것이다.



(증명). $\vec{u} = \mathrm{Proj}_U\vec{p}$라 하자. $\vec{w} \in U$이면 $\vec{u} - \vec{w} \in U$이므로 $\|\vec{p} - \vec{w}\|^2$은 다음과 같이 계산된다. ((2.5)에서 $\vec{w}$로 $\vec{u} - \vec{w}$를 취하면 $(\vec{p} - \vec{u}) \cdot (\vec{u} - \vec{w}) = 0$임.)

$$\begin{aligned}\|\vec{p} - \vec{w}\|^2 &= \|(\vec{p} - \vec{u}) + (\vec{u} - \vec{w})\|^2 \\ &= \|\vec{p} - \vec{u}\|^2 + \|\vec{u} - \vec{w}\|^2 + 2(\vec{p} - \vec{u}) \cdot (\vec{u} - \vec{w}) \\ &= \|\vec{p} - \vec{u}\|^2 + \|\vec{u} - \vec{w}\|^2 \\ &\geq \|\vec{p} - \vec{u}\|^2 .\end{aligned}$$

역으로 $\vec{u} \in U$가 $\vec{u} \neq \mathrm{Proj}_U\vec{p}$라 하자. 이 $\vec{u}$는 (2.6)을 만족하지 않음을 보이겠다. 즉 어떤 $\vec{w} \in U$에 대하여 $\|\vec{p} - \vec{w}\| < \|\vec{p} - \vec{u}\|$임을 보이도록 한다. 우리는 간단히 $\vec{w} = \mathrm{Proj}_U\vec{p}$로 둔다. 그러면 $\vec{w} \neq \vec{u}$로부터 $\|\vec{w} - \vec{u}\|^2 > 0$이므로

$$\begin{aligned}\|\vec{p} - \vec{u}\|^2 &= \|\vec{p} - \vec{w}\|^2 + \|\vec{w} - \vec{u}\|^2 + 2(\vec{p} - \vec{w}) \cdot (\vec{w} - \vec{u}) \\ &= \|\vec{p} - \vec{w}\|^2 + \|\vec{w} - \vec{u}\|^2 \\ &> \|\vec{p} - \vec{w}\|^2\end{aligned}$$

를 얻는다. □

이제 $\mathrm{Proj}_U\vec{p}$의 존재성을 증명해 보자. U가 1차원 공간일 때, 즉 $U = \mathrm{Span}(\vec{u}_1)$, $\vec{u}_1 \neq \vec{0}$ 형태일 때는 $\mathrm{Proj}_U\vec{p} = \frac{\vec{u}_1 \cdot \vec{p}}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1$으로 두면 된다. 왜냐하면 임의의 $\vec{w} \in U$는 적당한 스칼라 c에 대하여 $\vec{w} = c\vec{u}_1$으로 둘 수 있으므로

$$\left(\vec{p} - \frac{\vec{u}_1 \cdot \vec{p}}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1\right) \cdot (c\vec{u}_1) = c\,(\vec{p} \cdot \vec{u}_1 - \vec{u}_1 \cdot \vec{p}) = 0$$

가 되기 때문이다.

U가 2차원 공간일 때, 즉 $U = \mathrm{Span}(\vec{u}_1, \vec{u}_2)$ 형태일 때는

$$\mathrm{Proj}_U\vec{p} \overset{\mathrm{def}}{=} \mathrm{Proj}_{\vec{u}_1}\vec{p} + \mathrm{Proj}_{\vec{u}_2}\vec{p} \tag{2.7}$$

를 시도해 보자. 이때는 $\vec{w} = c_1\vec{u}_1 + c_2\vec{u}_2$로 둘 수 있으므로,

$$\mathrm{Proj}_{\vec{u}_1}\vec{p} = \frac{\vec{u}_1 \cdot \vec{p}}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 \stackrel{\text{def}}{=} d_1\vec{u}_1,$$
$$\mathrm{Proj}_{\vec{u}_2}\vec{p} = \frac{\vec{u}_2 \cdot \vec{p}}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 \stackrel{\text{def}}{=} d_2\vec{u}_2$$

으로 놓으면

$$\begin{aligned}(\vec{p} - \mathrm{Proj}_U\vec{p}) \cdot \vec{w} &= (\vec{p} - \mathrm{Proj}_{\vec{u}_1}\vec{p} - \mathrm{Proj}_{\vec{u}_2}\vec{p}) \cdot (c_1\vec{u}_1 + c_2\vec{u}_2) \\ &= c_1(\vec{p} - d_1\vec{u}_1 - d_2\vec{u}_2) \cdot \vec{u}_1 + c_2(\vec{p} - d_1\vec{u}_1 - d_2\vec{u}_2) \cdot \vec{u}_2 \\ &= c_1(\vec{p} - d_1\vec{u}_1) \cdot \vec{u}_1 + c_2(\vec{p} - d_2\vec{u}_2) \cdot \vec{u}_2 - (c_1d_2 + c_2d_1)\vec{u}_1 \cdot \vec{u}_2 \\ &= -(c_1d_2 + c_2d_1)(\vec{u}_1 \cdot \vec{u}_2)\end{aligned}$$

를 얻는다. 그러므로 $\vec{u}_1 \cdot \vec{u}_2 = 0$만 성립해 준다면 $\mathrm{Proj}_U\vec{p}$는 (2.7)에 의해서 얻어진다.

보조정리 *2.41* U의 차원이 n일 때, U의 기저 $\vec{u}_1, \ldots, \vec{u}_n$의 모든 벡터들이 서로 수직이기만 하면

$$Proj_U\vec{p} \stackrel{\text{def}}{=} Proj_{\vec{u}_1}\vec{p} + \cdots + Proj_{\vec{u}_n}\vec{p}$$

로 놓아 $Proj_U\vec{p}$를 얻을 수 있다. ⊣

보조정리 *2.42* $\vec{p} - Proj_U\vec{p}$는 모든 $\vec{u} \in U$와 수직이다. (이것은 [연습문제 *2.37*]을 일반화한 것으로 볼 수 있다.)

(증명). $\vec{u} \in U$는 $\vec{u}_i$들의 선형조합이므로 각 $i = 1, \ldots, n$에 대하여 $\vec{p} - \mathrm{Proj}_U\vec{p}$가 $\vec{u}_i$와 수직임을 보이면 된다.

$$\begin{aligned}(\vec{p} - \mathrm{Proj}_U\vec{p}) \cdot \vec{u}_i &= (\vec{p} - \mathrm{Proj}_{\vec{u}_1}\vec{p} - \cdots - \mathrm{Proj}_{\vec{u}_n}\vec{p}) \cdot \vec{u}_i \\ &= \vec{p} \cdot \vec{u}_i - 0 - \cdots - 0 - \frac{\vec{u}_i \cdot \vec{p}}{\vec{u}_i \cdot \vec{u}_i}\vec{u}_i \cdot \vec{u}_i - 0 \cdots - 0 \\ &= \vec{p} \cdot \vec{u}_i - \vec{u}_i \cdot \vec{p} = 0\end{aligned}$$

이다. ⊣

결국 직교기저(*orthogonal basis*)의 존재가 관건이다. 벡터공간의 임의의 기저 $\{\vec{u}_1, \ldots, \vec{u}_n\}$으로부터 직교기저 $\{\vec{u}_1^*, \ldots, \vec{u}_n^*\}$을 얻을 수 있음이 알려져 있으며 이 과정을 그램-슈미트 과정(*Gram-Schmidt process*)이라고 한다. 한점집합 기저 $\{\vec{u}_1\}$은 직교기저인 것으로 본다.

모든 원소가 단위벡터인 직교기저를 정규직교기저(*orthonormal basis*)라고 한다.

연습문제 2.43 벡터공간 V, 벡터 $\vec{u}_1, \vec{u}_2, \vec{v}_1, \vec{v}_2 \in V$, 스칼라 $c \in \mathbb{R}$ 및 부분공간 $U \subseteq V$ 에 대하여 다음의 사실들을 증명하시오.

(1) $\mathrm{Proj}_{c\vec{u}_1}\vec{v}_1 = \mathrm{Proj}_{\vec{u}_1}\vec{v}_1$

(2) $\mathrm{Proj}_U(c\vec{v}_1) = c\mathrm{Proj}_U\vec{v}_1$

(3) $\mathrm{Proj}_U(\vec{v}_1 + \vec{v}_2) = \mathrm{Proj}_U\vec{v}_1 + \mathrm{Proj}_U\vec{v}_2$

(4) $\mathrm{Proj}_U\vec{u} = \vec{u}$ for all $\vec{u} \in U$. ⊣

연습문제 2.44 V가 벡터공간이고 $\vec{p}_1, \ldots, \vec{p}_n \in V$, $U \subseteq V$는 부분공간일 때 아래의 식이 성립함을 증명하시오.[†]

$$\mathrm{Proj}_U\big(\mathrm{Span}(\vec{p}_1, \ldots, \vec{p}_n)\big) = \mathrm{Span}(\mathrm{Proj}_U\vec{p}_1, \ldots, \mathrm{Proj}_U\vec{p}_n)$$ ⊣

정리 *2.45* (*Gram-Schmidt Process*) 모든 내적공간은 정규직교기저를 가진다.

(증명). 유한차원 벡터공간에 대해서만 증명하겠다. 주어진 벡터공간을 U라 하고 이것의 차원을 n이라 하자. [정리 2.19]에 의하여 존재가 보장된 U의 기저 하나를 $\{\vec{u}_1, \ldots, \vec{u}_n\}$이라 하자. 각 $i = 1, \ldots, n$에 대하여 $U_i \overset{\text{def}}{=} \mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_i)$라 한다. 이제 다음과 같이 $\vec{u}_1^*, \ldots, \vec{u}_n^*$을 구성한다.

$$\begin{aligned}
\vec{u}_1^* &\overset{\text{def}}{=} \vec{u}_1, \\
\vec{u}_2^* &\overset{\text{def}}{=} \vec{u}_2 - \mathrm{Proj}_{U_1}\vec{u}_2, \\
&\ \vdots \\
\vec{u}_n^* &\overset{\text{def}}{=} \vec{u}_n - \mathrm{Proj}_{U_{n-1}}u_n.
\end{aligned}$$

그리고 각 $i = 1, \ldots, n$에 대하여 다음 사실을 수학적 귀납법에 의하여 증명한다.

(1) $\mathrm{Span}(\vec{u}_1^*, \ldots, \vec{u}_i^*) = U_i$

(2) $\vec{u}_1^*, \ldots, \vec{u}_i^*$의 임의의 두 벡터는 서로 수직이다.

다음 사실의 증명은 독자에게 연습문제로 남긴다. $\{\vec{u}_1^*, \ldots, \vec{u}_n^*\}$은 U의 직교기저이다. 그리고 $\{\vec{u}_1^*/\|\vec{u}_1^*\|, \ldots, \vec{u}_n^*/\|\vec{u}_n^*\|\}$은 우리가 원하던 정규직교기저이다. □

단평 2.46 직교기저를 얻는 방법은 그램-슈미트 과정 외에도 여러 가지가 있다. 대표적인 방법들로 Householder transformation, Givens rotation, Cholesky decomposition 등이 있다. 각 방법은 다른 방법들과 비교하여 장점과 단점을 가진다. ⊣

[†] $W = \{\vec{p}_1, \ldots, \vec{p}_n\}$로 두면 $\mathrm{Proj}_U(\mathrm{Span}(W)) = \mathrm{Span}(\mathrm{Proj}_U W)$로 쓸 수 있다. 즉 Span과 Proj 간에 교환법칙이 성립한다는 뜻.

단평 2.47 정규직교기저는 주어진 벡터를 이 기저의 선형조합으로 나타내기가 용이하다는 장점이 있다. 벡터공간 V의 정규직교기저 $\{\vec{u}_1, \ldots, \vec{u}_n\}$과 하나의 벡터 $\vec{v} \in V$가 주어졌을 때

$$\vec{v} = c_1\vec{u}_1 + \cdots + c_n\vec{u}_n \tag{2.8}$$

가 되는 계수 c_i들은 간단히 $c_i = \vec{u}_i \cdot \vec{v}$로 얻어진다. 이 사실은 (2.8)의 양변에 $\vec{u}_i$를 내적시키면 확인할 수 있다.

$\{\vec{u}_1, \ldots, \vec{u}_n\}$가 정규직교기저가 아닌 직교기저일 때는 $c_i = \frac{\vec{u}_i \cdot \vec{v}}{\vec{u}_i \cdot \vec{u}_i}$, 즉 사영으로 얻은 벡터의 계수임을 쉽게 알 수 있다. ⊣

연습문제 2.48 아래에 주어진 부분공간의 정규직교기저를 각각 구하고, 벡터 $\vec{v}_1$을 이 정규직교기저의 선형조합으로 나타내어라. 또한 $\vec{v}_2$를 이 부분공간에 사영하여라.

(1) $\{(x, y, z) \mid 2x + y - z = 0\} \subseteq \mathbb{R}^3$, $\vec{v}_1 = (3, -5, 1)$, $\vec{v}_2 = (2, 1, 3)$. 이 문제에 대해서는 정규직교기저를 2개 구해 보아라.

(2) $\mathrm{Span}((1, 0, 1, 1), (3, 1, 0, 0), (2, 1, 0, -2)) \subseteq \mathbb{R}^4$,
$\vec{v}_1 = (6, 2, 1, -1)$, $\vec{v}_2 = (0, 7, 1, -1)$. ⊣

연습문제 2.49 $U := \mathrm{Span}((1, 0, 1, 1), (3, 1, 0, 0), (2, 1, 0, -2)) \subseteq \mathbb{R}^4$로 두었을 때 $U = \{(x, y, z, w) \in \mathbb{R}^4 \mid ax + by + cz + dw = 0\}$가 되는 $a, b, c, d \in \mathbb{R}$을 구하여라. (힌트). U의 직교기저를 확장하여 $\mathbb{R}^4$의 직교기저를 얻으면 이때 직교기저에 새로 추가된 벡터가 곧 (a, b, c, d)가 된다. ⊣

연습문제 2.50 이제 직교기저를 이용하여 [연습문제 2.33]의 (ii)를 풀 수 있을 것이다. 이 문제에 등장하는 모든 벡터는 영벡터가 아니라고 가정한다.

3차원 공간에 선형독립인 두 벡터 $\vec{c}_1$, $\vec{c}_2$가 있고 $\vec{a} \cdot \vec{c}_i = \vec{b} \cdot \vec{c}_i = 0,\ (i = 1, 2)$일 때 $\vec{a} = k\vec{b}$인 스칼라 k가 존재한다. ⊣

연습문제 2.51 (Bessel inequality) n차원 벡터공간 V의 서로 다른 $m \leq n$개의 원소 $\vec{u}_1, \ldots, \vec{u}_m$이 주어졌으며, 이들 중 어떤 2개를 취해도 서로 수직이라고 한다. 임의의 $\vec{v} \in V$와 각 $i = 1, \ldots, m$에 대해서 $c_i \overset{\text{def}}{=} \vec{u}_i \cdot \vec{v}/\vec{u}_i \cdot \vec{u}_i$로 놓으면 모든 선형조합 $d_1\vec{u}_1 + \cdots + d_m\vec{u}_m$에 대해서 부등식

$$\|\vec{v} - \textstyle\sum_{i=1}^m c_i\vec{u}_i\| \leq \|\vec{v} - \sum_{i=1}^m d_i\vec{u}_i\|$$

가 성립함을 증명하여라. (힌트). [보조정리 2.40]을 이용한다. ⊣

연습문제 2.52 벡터공간 V의 부분집합 $B \overset{\text{def}}{=} \{\vec{u}_1, \ldots, \vec{u}_n\}$의 원소들은 모두 단위벡터이고 서로 수직이라고 한다. 만일 모든 $\vec{v} \in V$에 대해서

$$\|\vec{v}\|^2 = \textstyle\sum_{i=1}^n (\vec{u}_i \cdot \vec{v})^2$$

가 성립한다면 B는 V의 정규직교기저임을 보여라. (역은 당연히 참이다.)

(힌트). $\vec{v}' := \sum_{i=1}^{n}(\vec{u}_i \cdot \vec{v})\vec{u}_i$로 놓고 $\|\vec{v}' - \vec{v}\|^2 = 0$임을 보인다. ⊣

연습문제 2.53 $0 \le x \le 1$에서 정의된 연속함수들의 집합 다음과 같이 연산을 정의했을 때 $\mathbb{R}$을 스칼라로 하는 내적공간이 된다. (이것은 무한차원 벡터공간이다.)

(1) $(f+g)(x) = f(x) + g(x)$

(2) $(af)(x) = a(f(x))$

(3) $f \cdot g = \int_0^1 f(x)g(x)\,\mathrm{d}x$

$f_1(x) = x$, $f_2(x) = x^2$으로 두었을 때 $\mathrm{Span}(f_1, f_2)$의 정규직교기저를 구하여라. ⊣

2.3 직교여공간

정의 2.54 V를 내적공간, $U, W \subseteq V$를 부분집합이라 하자.

$\vec{v}_1, \vec{v}_2 \in V$는 $\vec{v}_1 \cdot \vec{v}_2 = 0$일 때 서로 수직이라고 하고 $\vec{v}_1 \perp \vec{v}_2$로 나타낸다.

$(\forall \vec{u} \in U)(\vec{u} \perp \vec{u})$일 때 $\vec{v}$는 U와 수직이라고 하고 $\vec{v} \perp U$로 나타낸다.

$(\forall \vec{w} \in W)(\vec{w} \perp U)$일 때 W와 U는 서로 수직이라고 하고 $W \perp U$로 나타낸다. 이것은 $U \perp W$와 동등하다.

U와 수직인 모든 벡터들의 집합을 $U^{\perp}$로 나타낸다. ⊣

단평 2.55 위 정의의 표기법을 따르면 [보조정리 2.42]는 $(\vec{p} - \mathrm{Proj}_U \vec{p}) \perp U$로 쓸 수 있다. 따라서 $\{\vec{u}_1, \ldots, \vec{u}_n\}$이 벡터부분공간 U의 직교기저이고 $\vec{p} \notin U$일 때

$$\{\vec{u}_1, \ldots, \vec{u}_n, \vec{p} - \mathrm{Proj}_U \vec{p}\}$$

는 $\mathrm{Span}(U \cup \{\vec{p}\})$의 직교기저가 된다. ⊣

연습문제 2.56 내적공간 V의 임의의 부분집합 U에 대해서 $U^{\perp}$는 벡터합과 스칼라곱에 대하여 닫혀있음을 증명하여라. (즉 부분공간임을 증명하여라.) ⊣

정의 2.57 두 부분공간 U와 W의 합 $U + W$는 다음과 같이 정의된다.

$$U + W \overset{\mathrm{def}}{=} \{\vec{u} + \vec{v} \mid \vec{u} \in U,\ \vec{w} \in W\}$$

부분공간 $U \subseteq V$의 V에 대한 직교여공간(*orthogonal complement*)은 다음과 같은 조건을 만족하는 부분공간 $W \subseteq V$를 뜻한다. 직교여공간은 수직여공간이라고 하기도 한다.

(*i*) $U + W = V$

(*ii*) $U \cap W = \{\vec{0}\}$

(iii) $U \perp W$

W가 위의 3 조건들 중 (i)와 (ii) 2개만 만족하는 경우 W는 U의 여공간(*complement*)이라고 한다. (W가 U의 (직교)여공간인 것은 U가 W의 (직교)여공간일 필요충분조건이다.) $\dashv$

<u>연습문제</u> 2.58 W와 U가 V의 부분공간이면 $W + U \subsetneq V$인 경우에도 $W + U$는 V의 부분공간임을 증명하여라. $\dashv$

<u>연습문제</u> 2.59 W와 U와 V 안에서 서로 여공간이라는 것은 V의 모든 원소 $\vec{v}$에 대해서 $\vec{u} \in U$, $\vec{w} \in W$, $\vec{u} + \vec{w} = \vec{v}$인 $\vec{u}$와 $\vec{w}$가 유일하게 존재한다는 것의 필요충분조건임을 증명하여라. (힌트). $\vec{u} \in U$, $\vec{w} \in W$이고 $\vec{u} + \vec{w} = \vec{0}$이면 $\vec{u} = \vec{w} = \vec{0}$임을 보인다. $\dashv$

<u>연습문제</u> 2.60 $U \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}^2 \mid 2x + y = 0\}$의 $\mathbb{R}^2$ 안에서의 여공간을 3개 구하시오. 이들 중 하나는 직교여공간이어야 한다. $\dashv$

<u>정리</u> *2.61* 유한차원 내적공간 V와 부분공간 $U \subseteq V$가 주어졌을 때

$$\dim(V) = \dim(U) + \dim(U^\perp)$$

이다.

(증명). U의 정규직교기저를 $\{\vec{u}_1, \ldots, \vec{u}_n\}$라 두고 이것을 그램-슈미트 과정을 통하여 확장하여 V의 정규직교기저 $\{\vec{u}_1, \ldots, \vec{u}_{n+m}\}$을 얻은 다음, $\{\vec{u}_{n+1}, \ldots, \vec{u}_{n+m}\}$이 $U^\perp$의 (정규직교)기저임을 다음과 같이 보이면 된다.

$\mathrm{Span}(\vec{u}_{n+1}, \ldots, \vec{u}_{n+m}) \subseteq U^\perp$ 임은 당연하다. $\supseteq$를 보이기 위하여 $\vec{v} \in U^\perp$라 하자. $\vec{v} \in V$이므로 $\vec{v} = c_1\vec{u}_1 + \cdots + c_{n+m}\vec{u}_{n+m}$으로 나타낼 수 있다. $i = 1, \ldots, n$에 대하여

$$\vec{v} \cdot \vec{u}_i = 0 = c_i \|\vec{u}_i\|^2$$

이므로 $c_i = 0$이고, 따라서 $\vec{v} = c_{n+1}\vec{u}_{n+1} + \cdots + c_{n+m}\vec{u}_{n+m}$이 됨을 알 수 있다. $\square$

<u>연습문제</u> 2.62 내적공간 V와 부분공간 $U \subseteq V$가 주어졌을 때, U의 직교여공간은 유일하게 존재하며 그것은 $U^\perp$임을 증명하여라. $\dashv$

<u>연습문제</u> 2.63 V가 벡터공간이고 $U \subseteq V$일 때 $\left(U^\perp\right)^\perp = \mathrm{Span}(U)$임을 증명하여라. $\dashv$

3
행렬

<u>정의</u> 3.1 X가 임의의 공아닌 집합이고 m, n이 양의 정수일 때 X 위의 $m\times n$ 행렬(*matrix*)이란 길이 n인 m개의 행으로 이루어진, X의 원소들의 '2차원적 배열'을 말한다. 집합론적으로 엄격하게 말하자면 $m \times n$에서 X로 가는 함수가 곧 X 위의 $m \times n$ 행렬이다. 그러니까 X는 행렬의 공역(*codomain*)이라고 말할 수 있다.

X 위의 모든 $m \times n$ 행렬들의 집합을 $X^{m\times n}$으로 나타낸다.

행(*row*), 열(*column*), 행벡터(*row vector*), 열벡터(*column vector*)의 의미는 잘 알려져 있으므로 정의 없이 사용하기로 한다.

$m = n$일 때 이 행렬을 정방행렬(*square matrix*)이라고 부른다. ⊣

행렬의 공역 X로 흔히 쓰이는 집합으로는 $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$, various functions, finite fields 등이 있다. 이 과목에서는 거의 대부분의 경우 $\mathbb{R}$을 공역으로 하는 행렬을 다룰 것이며, 그렇지 않은 경우에는 이 점을 명시할 것이다.

행렬의 i-행, j-열에 있는 X의 원소를 ij-항(*entry*), 또는 ij-원소(*element*)라고 부른다. 행렬은 보통 대문자 A, B 등으로 표기하며 A의 ij-항은 보통 $(A)_{ij}$ 혹은 a_{ij}로 나타낸다.

각 $m, n > 0$에 대해서 $m \times n$ 행렬들은 mn-차원 벡터공간을 이룬다. 이때 벡터합과 은 당연한 방법으로 정의한다.

3.1 행렬곱

행렬합과 스칼라곱은 당연한 방법으로 정의하지만 행렬곱의 정의는 썩 당연하지 않게 보일 것이다.

<u>정의</u> 3.2 A가 $m\times r$ 행렬이고 B가 $r\times n$ 행렬일 때 이 둘의 행렬곱(*matrix multiplication*) AB는

(1) Size가 $m \times n$이고

(2) ij-항은 $a_{i1}b_{1j} + \cdots + a_{ir}b_{rj}$ 인

행렬을 뜻한다. $\dashv$

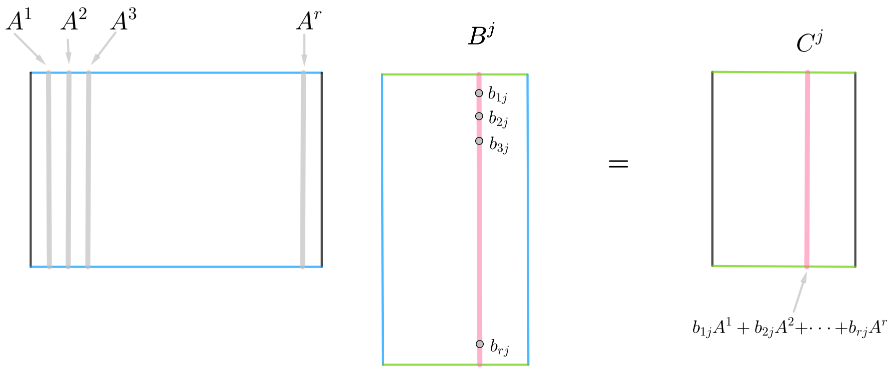
<u>단평</u> 3.3 행렬곱 AB는 A의 열의 개수가 B의 행의 개수와 일치할 때만 정의된다. A와 B가 정방행렬일 때는 이 둘의 크기가 같을 때 행렬곱이 정의된다. 정방행렬이 아닐 때는 크기가 $m \times n$으로 같은 두 행렬의 곱은 정의되지 않는다.

행렬곱 $C = AB$를 보는 3가지 방법이 있으며 이들을 아래의 (1), (2), (3)에 각각 보였다. 행렬 M의 i-행을 M_i, j-열을 M^j로 나타내기로 한다.

(1) $(C)_{ij}$는 A_i와 B^j의 내적이다.

(2) C^j는 A와 B^j의 곱이며 B의 다른 열들과는 하등 상관이 없다. C^j는 A의 열벡터들의 선형조합으로서 B^j의 항들을 계수로 가진다.

$$A \in \mathbb{R}^{m\times r}, \quad B \in \mathbb{R}^{r\times n}, \quad C \in \mathbb{R}^{m\times n}$$



(3) C_i는 A_i와 B의 곱이며 A의 다른 행들과는 하등 상관이 없다. C_i는 B의 행벡터들의 선형조합으로 A_i의 항들을 계수로 가진다. $\dashv$

$\mathbb{R}^n$의 원소, 즉 벡터를 행렬로 볼 때는 $\mathbb{R}^{1\times n}$의 원소로 볼 수도 있고 $\mathbb{R}^{n\times 1}$의 원소로 볼 수도 있을 것인데, 우리는 항상 후자, 즉 열벡터로 보기로 한다. 따라서 벡터간의 내적은 행렬곱을 이용하여 다음과 같이 나타낸다.

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} \tag{3.1}$$

물론 이때 벡터 내적의 값인 스칼라는 1×1-행렬과 동일시 한다.

<u>보조정리</u> *3.4* 행렬곱은 결합법칙과 분배법칙은 만족하지만 교환법칙은 만족하지 않는다.

(증명). 교환법칙을 만족하지 않는 것을 보이기는 아주 쉽다. 아래에 한 예를 들었다.

$$\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}0 & 3\\ 5 & 0\end{bmatrix} = \begin{bmatrix}0 & 3\\ 10 & 0\end{bmatrix} \neq \begin{bmatrix}0 & 6\\ 5 & 0\end{bmatrix} = \begin{bmatrix}0 & 3\\ 5 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}$$

결합법칙을 만족하는 것을 보이는 것은, 수학적 내용은 별로 여러울 것 없지만 표기법이 약간 까다롭다고 생각될 수 있다. $A \in \mathbb{R}^{n\times r}$, $B \in \mathbb{R}^{r\times s}$, $C \in \mathbb{R}^{s\times m}$가 주어졌을 때 각 $i = 1, \ldots n$, $j = 1, \ldots, m$에 대해서

$$((AB)C)_{ij} = (A(BC))_{ij}$$

를 보이면 된다. 위 식의 좌변은

$$\sum_{k=1}^{s}(AB)_{ik}(C)_{kj} = \sum_{k=1}^{s}\Big(\sum_{\ell=1}^{r} a_{i\ell}b_{\ell k}\Big)c_{kj} = \mathrm{Sum}\left\{a_{i\ell}b_{\ell k}c_{kj} \mid \ell \le r,\ k \le s\right\}$$

이고 우변은

$$\sum_{\ell=1}^{r}(A)_{i\ell}(BC)_{\ell j} = \sum_{\ell=1}^{r} a_{i\ell}\Big(\sum_{k=1}^{s} b_{\ell k}c_{kj}\Big) = \mathrm{Sum}\left\{a_{i\ell}b_{\ell k}c_{kj} \mid \ell \le r,\ k \le s\right\}$$

이므로 값이 서로 같다.

분배법칙에 대한 증명은 생략한다. □

<u>단평</u> 3.5 행렬곱을 손으로 계산하면 시간이 많이 걸리고 계산 실수도 많이 발생한다. 행렬곱을 포함한 행렬 관련 다양한 계산을 위한 소프트웨어는 많이 있지만 그 중 간편하고 무료인 것으로 `https://www.wolframalpha.com`과 `GeoGebra`를 추천한다. ⊣

<u>정의</u> 3.6 정방행렬 중에서

$$\text{if } i = j \text{ then } a_{ij} = 1 \text{ else } a_{ij} = 0$$

인 A를 단위행렬(*identity matrix*)라고 부르고 I로 나타낸다. 때로는 I_n과 같이 I의 행의 개수(=열의 개수) n을 첨자로 사용하기도 한다.

정방행렬 A가 주어졌을 때

$$AB = BA = I$$

인 정방행렬 B를 A의 역행렬(*inverse matrix*)이라고 부르고 A^{-1}로 나타낸다. ⊣

정의 3.7 역행렬이 존재하는 행렬을 가역행렬(*invertible matrix*)이라고 부른다. ⊣

단위행렬과 역행렬의 기본성질은 다음과 같다.

(1) 단위행렬은 기본벡터들을 열벡터로(혹은 행벡터로) 가지는 행렬이라고 말할 수 있다.

(2) $A \in \mathbb{R}^{m\times n}$이면 $AI_n = A = I_m A$이다.

(3) A, B, C는 행렬이고 이 중에 A는 가역행렬이라 하자. $AB = C$이면 $B = A^{-1}C$이고, $BA = C$이면 $B = CA^{-1}$이다.

연습문제 3.8 어떤 행렬의 역행렬이 존재하면 그것은 유일함을 보이시오. 역행렬이 존재하는 행렬과 존재하지 않는 행렬의 예를 하나씩 들어라. 가역행렬 A의 역행렬 A^{-1}는 가역행렬이며 A^{-1}의 역행렬은 A임을 보여라.

가역행렬 A와 B의 곱 AB는 가역행렬이며 그 역행렬은 $B^{-1}A^{-1}$ 임을 보여라. ⊣

단평 3.9 $AB = I$ 이면 $BA = I$이다. 이 사실은 그리 쉽게 증명되지 않는다. 이를 증명하는 데는 행렬의 "랭크" 개념을 도입해야 하며, 이 개념은 뒤에 공부할 "선형사상"의 이론을 통해 이해할 수 있다. (See p59.) ⊣

연습문제 3.10 행렬의 모든 항이 0인 행렬을 영행렬(*zero matrix*)이라고 하고 O로 나타낸다.

다음의 함의명제들은 모두 틀렸음을 보여라.

(1) $AB = O \;\Rightarrow\; A = O$ or $B = 0$

(2) $A^2 = O \;\Rightarrow\; A = O$

(i)만 보이면 (ii)는 자동으로 증명된 셈인가? (ii)만 보이면 (i)는 자동으로 증명된 셈인가?

정방행렬 A, B에 대하여 다음의 명제들을 증명하여라.

(1) $AB = O \;\Rightarrow\; A^{-1}$이 존재하지 않음 or B^{-1}이 존재하지 않음

(2) $A^2 = O \;\Rightarrow\; A^{-1}$이 존재하지 않음

(i)만 보이면 (ii)는 자동으로 증명된 셈인가? (ii)만 보이면 (i)는 자동으로 증명된 셈인가? ⊣

<u>정의</u> 3.11 정방행렬 A가

$$a_{ij} = 0 \text{ for all } i \neq j$$

를 만족하면 대각행렬(*diagonal matrix*)이라고 부르며,

$$a_{ij} = 0 \text{ for all } i > j$$

를 만족하면 상삼각행렬(*upper triangular matrix*),

$$a_{ij} = 0 \text{ for all } i < j$$

를 만족하면 하삼각행렬(*lower triangular matrix*)라고 부른다. ⊣

<u>연습문제</u> 3.12 다음 사실들을 증명하여라: 하삼각행렬들의 곱은 하삼각형행렬이고 상삼각행렬들의 곱은 상삼각형행렬이다. 대각행렬들의 곱은 대각행렬이 되며, 이때는 교환법칙이 성립한다. 일반적인 행렬과 대각행렬의 곱에서는 교환법칙이 성립하지 않는다. ⊣

<u>연습문제</u> 3.13 정방대각행렬이 가역행렬일 필요충분조건을 말하고, 가역대각행렬의 역행렬을 구하는 방법을 설명하여라. ⊣

<u>정의</u> 3.14 A가 $m \times n$ 실행렬일 때 A의 전치행렬(*transpose*) A^T는

(1) 크기가 $n \times m$이고

(2) ij-항은 a_{ji}인

행렬을 뜻한다. 대칭행렬(*symmetric matrix*)이란 $A^T = A$인 행렬 A를 말한다. (대칭행렬은 당연히 정방행렬이어야만 한다.)

벡터의 내적 $\vec{x} \cdot \vec{y}$는 벡터들을 열벡터로 보고 행렬곱을 사용하여 $\vec{x}^T \vec{y}$로 나타낼 수 있다. ⊣

전치행렬의 정의는 복소행렬에도 그대로 적용될 수 있지만 실행렬에서만큼 의미가 크지 않다. 복소행렬에 대해서는 실행렬의 전치에 대응되는 에르미트 전치(*hermitian transpose*)라는 개념을 사용한다.

<u>정의</u> 3.15 A가 복소행렬일 때 A의 전치행렬을 취한 다음 각 항을 그것의 켤레복소수로 바꾸어 얻은 행렬을 에르미트 전치행렬(*hermitian transpose*)라고 하고 A^H로 나타낸다. 복소 벡터의 내적 $\vec{x} \cdot \vec{y}$는 벡터들을 열벡터로 보고 행렬곱을 사용하여 $\vec{x}^H \vec{y}$로 나타낼 수 있다.

에르미트 행렬(*Hermitian matrix*)란 $A^H = A$인 (정방)행렬 A를 말한다. ⊣

연습문제 3.16

$$(AB)^T = B^T A^T, \quad (AB)^H = B^H A^H, \quad (A^T)^T = A = (A^H)^H$$

임을 보이시오. 일반적으로 대칭행렬들의 곱은 대칭행렬이 아님을 보여라. ⊣

연습문제 3.17

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

임을 보여라. ⊣

3.2 고교수학에서 다루었던 행렬

고등학교 수학에서는 2×2-실행렬만 다룬다. 그리고 단위행렬은 I가 아니라 E로 나타낸다.

행렬식(디터미넌트, *determinant*)의 정의부터 시작하자. 디터미넌트는 69쪽에서 상세하게 다룰 것이다. 일단 여기서 2×2 행렬에 대한 이론만 알아 보자.

정의 3.18 행렬 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$의 "디터미넌트"는 실수 $ad - bc$를 뜻하며 기호로는 $\det(A)$, 혹은 $|A|$로 나타낸다. ⊣

다음의 사실들은 아주 쉽게 알 수 있다.

(1) $\det(E) = 1$

(2) $\det(O) = 0$

(3) 행렬의 두 행(혹은 두 열)이 동일하면 디터미넌트는 0이다.

(4) 행렬의 두 행(혹은 두 열)을 맞바꾸면 디터미넌트의 절댓값은 변하지 않고 부호만 바뀐다.

(5) $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$이면 $A^{-1} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$이다.

(6) $|A| \neq 0 \ \Leftrightarrow\ (A^{-1}\text{ 존재})$

보조정리 *3.19* 행렬 A의 하나의 행(혹은 열)이 다른 행(혹은 열)의 실수 배라는 것은, 즉 행벡터들이(혹은 열벡터들이) 선형종속이라는 것은 $\det(A) = 0$일 필요충분조건이다.

(증명). 충분조건임을 증명하기는 쉽다. 필요조건임을 증명하는 것은 그리 어려울 것은 없지만 약간의 주의를 요한다. □

<u>보조정리</u> *3.20* 행렬의 곱의 디터미넌트는 각 행렬의 디터미넌트의 곱과 같다. 즉,

$$\det(AB) = \det(A)\det(B).$$

(증명). 그냥 $A = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$, $B = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$로 놓고 계산하면 된다. □

2원1차 연립방정식

$$\begin{aligned} &ax + by = u, \\ &cx + dy = v \end{aligned}$$

는 행렬을 이용하여

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

로 나타낼 수 있다.

<u>정리</u> *3.21 2원1차* 연립방정식 $A\vec{X} = \vec{0}$가 $\vec{X} \neq \vec{0}$를 해로 가질 필요충분조건 두 개를 제시한다. ① $\det(A) = 0$, ② 해를 무한히 많이 가진다.

(증명). (①에 대해서만)

$\Rightarrow$: $\det(A) \neq 0$이면 A^{-1}을 연립방정식의 양변의 왼쪽에서 곱한다.

$\Leftarrow$: $\det(A) = ad - bc = 0$라고 하자. $\vec{X} = (x, y)^T$로 둔다.

(1) $a = 0$인 경우

① $c = 0$인 경우: $(x, y) = (1, 0)$은 $A\vec{X} = \vec{0}$의 0아닌 해이다.

② $c \neq 0$인 경우: $b = 0$이어야만 한다. $(x, y) = (-\frac{d}{c}, 1)$은 $A\vec{X} = \vec{0}$의 0아닌 해이다.

(2) $a \neq 0$인 경우: $d = \frac{bc}{a}$이므로 $(x, y) = (-\frac{b}{a}, 1)$는 $A\vec{X} = \vec{0}$의 0아닌 해이다. 확인 계산을 아래에 보였다.

$$A\vec{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -\frac{b}{a} \\ 1 \end{bmatrix} = \begin{bmatrix} -b + b \\ -\frac{bc}{a} + d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ □

<u>보조정리</u> *3.22* $A \in \mathbb{R}^{2\times 2}$일 때, 모든 $n \geq 3$에 대해서

$$A^n = O \iff A^2 = O.$$

(증명). $\Leftarrow$는 당연하므로 $\Rightarrow$만 보이겠다. $n = 3$인 경우에 대해서만 증명을 보이겠다. 일반적인 경우의 증명은 귀납법을 사용하면 된다.

$A^3 = O$를 가정한다. 그러면 $\det(A^3) = \det(O) = 0$가 될 것이다. 그런데 $\det(A^3) = \det(A)^3$이므로 이제 $\det(A) = 0$임을 알 수 있다.

A의 첫 행과 둘째 행이 모두 $(0, 0)$이라면 $A^2 = O$는 당연히 성립한다. 그러므로 둘 중의 적어도 하나가 $(0, 0)$이 아닌 경우에 대해서 $A^2 = O$를 증명하면 된다.

A의 첫 행이 $(0, 0)$이 아니라는 가정 하에 증명한다. 둘째 행이 $(0, 0)$이 아닐 때는 비슷한 방법으로 증명할 수 있다. 또한 $(A)_{11}$이 0이 아니라고 가정한다. $(A)_{12}$가 0이 아닐 때는 비슷한 방법으로 증명할 수 있다. 이제 정리 3.19를 이용하여

$$A = u \begin{bmatrix} 1 & a \\ k & ka \end{bmatrix} \tag{3.2}$$

로 놓는다. 그러면 다음과 같은 식을 얻을 수 있다.

$$A^2 = u^2 \begin{bmatrix} 1 + ak & a(1 + ak) \\ k(1 + ak) & ka(1 + ak) \end{bmatrix} = u^2(1 + ak) \begin{bmatrix} 1 & a \\ k & ka \end{bmatrix}, \tag{3.3}$$

$$A^3 = u^3(1 + ak)^2 \begin{bmatrix} 1 & a \\ k & ka \end{bmatrix}. \tag{3.4}$$

이하는 생략한다. □

<u>고찰</u> 3.23 $A \in \mathbb{R}^m$에 대해서 $(\forall n \geq 3)(A^n = O \iff A^2 = O)$는 $m = 2$일 때만 성립한다. $m = 3$일 때는 행렬 [011/001/000]이 반례가 된다. $m \geq 4$에 대해서 거듭제곱이 0-행렬이 되는 정방행렬들을 조사해 보라. ⊣

3.3 선형연립방정식과 가우스 소거

n개의 미지수 $x_1, \ldots, x_n$에 대한 선형연립방정식은 다음과 같은 형식을 취한다. 이러한 방정식의 해의 개수는 0, 1, 또는 무한이 된다.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \tag{3.5}$$

m은 방정식의 개수이며 n보다 작을 수도, 클 수도, 같을 수도 있다. 행렬 $A \in \mathbb{R}^{m\times n}$를 $(A)_{ij} = a_{ij}$로 정의하고, $B \in \mathbb{R}^{m\times 1}$을 $(B)_{i1} = b_i$로 정의하고 미지수들의 $n \times 1$ 행렬을 $X = (x_1, \ldots, x_n)^T$로 두면, 위의 방정식은 다음과 같이 간단히 나타낼 수 있다.

$$AX = B \tag{3.6}$$

X와 B는 단 하나의 열만 가지고 있으므로 열벡터로 볼 수 있다. 이들이 벡터임을 강조하기 위하여 다음과 같이 쓰기도 한다.

$$A\vec{x} = \vec{b} \tag{3.7}$$

선형연립방정식 (3.5)를 풀기 위한 방법으로 가장 대표적인 것이 가우스 소거 (*Gaussian elimination*)이다. 이제 곧 알게 되겠지만 가우스 소거법은 우리가 이미 잘 알고 있는 선형 연립방정식의 해법을 행렬곱을 이용하여 나타내는 것이라고 생각하면 된다.

가우스 소거는 특정한 행렬을 (3.6)의 양변에 계속 곱하는 과정이라고 할 수 있는데 이러한 행렬곱 과정은 (3.6)에서 B가 여러 개의 열벡터로 구성된 경우에도 (이때 X도 여러 개의 열벡터로 구성되어야 함은 당연하다.) 그대로 적용할 수 있으며 이를 가우스-조던 소거(*Gauss-Jordan Elimination*)라고 한다. 바꾸어 말하면 가우스 소거는 가우스-조던 소거에서 B의 열의 개수가 1인 특수한 경우로 보아도 된다.

<u>정의</u> 3.24 행렬에 대한 기본 행작업(*elementary row operation*)은 아래에 보인 3가지 유형 중 하나의 작업을 뜻한다.

(1) 두 행의 위치를 서로 뒤바꾼다.

(2) 하나의 행에 0아닌 스칼라를 곱한다.

(3) 행 하나를 택하여 스칼라곱을 한 것을 다른 행에 더하여 새로운 행렬을 얻는다.⊣

정의 3.25 기본 행작업 행렬(*elementary matrix*)은 단위행렬 I_m에 기본 행작업을 실행하여 얻은 행렬을 뜻한다.

(1) (타입 I) 각 $1 \leq i \neq j \leq n$에 대해서 P_{ij}는 I_m의 i번째 행과 j번째 행을 뒤바꾸어 얻은 행렬을 뜻한다.

(2) (타입 II) 각 $1 \leq i \leq m$, $0 \neq \alpha \in \mathbb{R}$에 대해서 $I_{m,i}(\alpha)$는 I_m에서 (i,i)항 단 하나만을 α로 바꾸어 얻은 행렬을 뜻한다.

(3) (타입 III) 각 $\alpha \in \mathbb{R}$과 $1 \leq i \neq j \leq n$에 대해서 $E_{ij}(\alpha) \in \mathbb{R}^{m\times m}$는

$$E_{ij}(\alpha) \overset{\text{def}}{=} \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ & \vdots & & \dots & \\ 0 & \alpha & & \dots & 0 \\ 0 & 0 & \dots & \dots & 1 \end{bmatrix} \quad (\alpha \text{ is the } ij-\text{entry})$$

이다. 그러니까 $E_{ij}(\alpha)$는 $m \times m$-행렬로서 각 $1 \leq r \leq m$, $1 \leq s \leq m$에 대해서 rs-항(entry)은 $r = s$이면 1이고, $(r,s) = (i,j)$이면 α이며 그 외의 경우는 모두 0이 되는 것이다. 더 쉽게 말하자면 I_m에서 j번째 행에 α를 곱한 것을 i번째 행에 더하여 얻은 행렬이다. ⊣

예제 3.26 기본 행작업들을 아래의 행렬에 행해 보자.

$$A \overset{\text{def}}{=} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ -1 & 3 & 2 \end{bmatrix}$$

다음의 3 작업을 차례로 A에 실행한다.

(i) 1행과 2행을 뒤바꾼다. (ii) 3행에 2를 곱한다. (iii) 1행에 2를 곱한 것을 3행에 더한다.

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ -1 & 3 & 2 \end{bmatrix} \overset{(i)}{\Rightarrow} \begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \\ -1 & 3 & 2 \end{bmatrix} \overset{(ii)}{\Rightarrow} \begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \\ -2 & 6 & 4 \end{bmatrix} \overset{(iii)}{\Rightarrow} \begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \\ -2 & 8 & 8 \end{bmatrix}$$

위의 작업들은 A에 다음 행렬들을 왼쪽에서 차례로 곱하는 것에 대응된다.

$$P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_{3,3}(2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad E_{31}(2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

이 예에서 행렬 $A \in \mathbb{R}^{3\times 3}$에 어떤 기본 행작업을 행하는 것은, 그 작업을 I_3에 행하여 얻은 (기본 행작업) 행렬을 A의 왼쪽에서 곱하는 것과 같음을 확인하였다. 이 확인과정에서 p40에 있는 Remark 3.3의 (3)번 항목이 유용할 것이다. 이 사실은 일반적인 $A \in \mathbb{R}^{m\times n}$과 I_m에 대하여 엄격하게 증명할 수 있으나 여기서는 생략하겠다. ⊣

기본 열작업 행렬도 생각해 볼 수 있다. 이것은 주어진 행렬의 오른쪽에서 곱하여 기본 열작업, 즉 하나의 열을 택하여 스칼라곱을 한 것을 다른 열에 더하여 새로운 행렬을 얻는 작업에 대응된다. 우리는 기본 열작업에 대한 공부는 생략한다.

<u>연습문제</u> 3.27 기본 행작업 행렬들은 모두 가역행렬임을 보여라. 따라서 정방행렬의 가역성은 기본 행작업에 의하여 바뀌지 않는다. ⊣

이제 p47에서 말했던 가우스 소거를 설명하기 위한 모든 준비작업이 완료되었다. 가우스 소거란 선형연립방정식 (3.7), 혹은 (3.5)의 해를 구하는 체계적 방법이라고 했었음을 상기하라. 가우스 소거는 계수행렬(coefficient matrix) A가 정방행렬이 아닐 때도 행할 수 있지만 우리는 복잡한 상황을 피하기 위하여 당분간, §3.3의 끝까지는, A가 정방행렬임을, 즉 $m = n$임을 가정하고 논의를 진행할 것이다.

표기상의 편의를 위하여 선형연립방정식 $AX = B$에서 A의 오른쪽에 B를 더하여 얻은 행렬을 $[A|B]$로 나타내기로 한다.

아래의 예에서 시작하자.

$$\begin{array}{rrrll} x_1 & -x_2 & +2x_3 & = 3 & ① \\ 2x_1 & & +2x_3 & = 6 & ② \\ & x_2 & +3x_3 & = 4 & ③ \end{array}$$

이 연립방정식을 통상적인 방법으로 풀 때 첫 단계는 ②에서 ①의 2배를 빼는 것이다. 그 다음은 새롭게 얻은 ②를 2로 나누는 것이다. 계수들로 이루어진 행렬을 가지고 작업한다면 다음과 같이 진행할 수 있다.

$$[A|B] = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 2 & 0 & 2 & 6 \\ 0 & 1 & 3 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 2 & -2 & 0 \\ 0 & 1 & 3 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 3 & 4 \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix}1 & -1 & 2 & 3\\ 0 & 1 & -1 & 0\\ 0 & 0 & 4 & 4\end{bmatrix} \Rightarrow \begin{bmatrix}1 & -1 & 2 & 3\\ 0 & 1 & -1 & 0\\ 0 & 0 & 1 & 1\end{bmatrix} \Rightarrow \begin{bmatrix}1 & 0 & 1 & 3\\ 0 & 1 & -1 & 0\\ 0 & 0 & 1 & 1\end{bmatrix} \Rightarrow$$

$$\begin{bmatrix}1 & 0 & 0 & 2\\ 0 & 1 & -1 & 0\\ 0 & 0 & 1 & 1\end{bmatrix} \Rightarrow \begin{bmatrix}1 & 0 & 0 & 2\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\end{bmatrix}$$

마지막 행렬은 아래의 연립방정식에 대응된다.

$$\begin{aligned} x_1 \qquad\qquad &= 2 \quad ① \\ x_2 \qquad &= 1 \quad ② \\ x_3 &= 1 \quad ③ \end{aligned}$$

결국 다음과 같은 해를 얻었다. $(x_1, x_2, x_3) = (2, 1, 1)$.

이 과정에서 실행한 작업들은 모두 기본적 행작업이며 최종 결과는 다음과 같은 행렬을 $[A|B] \in \mathbb{R}^{3\times 4}$의 왼쪽에 곱하여 얻을 수 있다.

$$E_{23}(1)\ E_{13}(-1)\ E_{12}(1)\ I_{3,3}(\tfrac{1}{4})\ E_{32}(-1)\ I_{3,2}(\tfrac{1}{2})\ E_{21}(-2)$$

위에 보인 7개의 기본 행작업 행렬은 모두 타입이 II와 III이며 타입 I은 나타나지 않았다. 이것은 운이 좋았기 때문이라고 말할 수 있다. 타입 I 행작업이 처음부터 필요한 예로 $(A)_{11} = 0$인 경우를 들 수 있다. 타입 I 행작업이 두번째 단계에서 필요한 예로

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

를 들 수 있다. 이 예에서는 처음에 $E_{21}(-2)$를 곱하고 이어서 P_{23}을 곱해야 할 것이다.

———— ∘ ———— ∘ ————

이제 일반적인 선형연립방정식 $AX = B$, 즉 (3.6)의 해를 구하는 과정을 기술하겠다. (3.6)의 양변에 $n-1$개의 기본 행작업 행렬

$$E_{21}\left(-\frac{a_{21}}{a_{11}}\right),\ E_{31}\left(-\frac{a_{31}}{a_{11}}\right), \ldots,\ E_{n1}\left(-\frac{a_{n1}}{a_{11}}\right)$$

을 차례로 왼쪽에서 곱한다. 여기서 $a_{11} \neq 0$이 필요하다. 만일 $a_{11} = 0$이라면 $a_{k1} \neq 0$

인 k를 찾아서 P_{1k}를 곱해서 해결한다. 만일 이러한 k가 없다면 이때의 계수행렬 A는 가우스 소거를 받아들이지 않는다. (즉, 가우스 소거는 실패한다.) 그리고 이 연립방정식의 해는 존재하지 않거나 무한히 많다. (이에 대한 설명은 뒤에 공부할 것임.)

그러면 이제 (3.6)은 아래와 같이 바뀔 것이다. ($\frac{a_{i1}}{a_{11}}$는 생략하였다.)

$$E_{n1}E_{n-1,1}\cdots E_{21}AX = E_{n1}E_{n-1,1}\cdots E_{21}B. \tag{3.8}$$

A를 변형시켜 얻은 $E_{n1}E_{n-1,1}\cdots E_{21}A \overset{\text{def}}{=} A'$은 첫번째 열의 (1,1)-항 아래의 모든 항이 0임을 주목하라.

$E_{n1}E_{n-1,1}\cdots E_{21}B \overset{\text{def}}{=} B'$으로 놓으면 (3.8)은 다음과 같이 쓸 수 있다.

$$A'X = B' \tag{3.9}$$

다음에는 A'의 두번째 행에 대해서 작업한다.

(3.8)의 양변에 $n-2$개의 기본 행작업 행렬

$$E_{22}\left(-\frac{a'_{22}}{a'_{22}}\right),\ E_{32}\left(-\frac{a'_{32}}{a'_{22}}\right),\ldots,\ E_{n2}\left(-\frac{a'_{n2}}{a'_{22}}\right)$$

을 차례로 왼쪽에서 곱한다. 여기서 $a'_{22}\neq 0$이 필요하다. 만일 $a'_{22}=0$이라면 $a'_{k2}\neq 0$인 k를 찾아서 P_{2k}를 곱해서 해결한다. 만일 이러한 k가 없다면 계수행렬 A는 가우스 소거를 받아들이지 않는다. 그리고 이 연립방정식의 해는 존재하지 않거나 무한히 많다. 그러면 이제 (3.8)은 아래와 같이 바뀔 것이다. ($\frac{a'_{i2}}{a'_{22}}$는 생략하였다.)

$$E_{n2}E_{n-1,2}\cdots E_{32}A'X = E_{n2}E_{n-1,2}\cdots E_{32}B'. \tag{3.10}$$

A'을 변형시켜 얻은 $E_{n2}E_{n-1,2}\cdots E_{31}A' \overset{\text{def}}{=} A''$은 두번째 열의 (2,2)-항 아래의 모든 항이 0임을 주목하라.

$E_{n2}E_{n-1,2}\cdots E_{32}B' \overset{\text{def}}{=} B''$으로 놓으면 (3.10)은 다음과 같이 쓸 수 있다.

$$A''X = B'' \tag{3.11}$$

이런 과정을 반복하면 결국 (3.6)은 다음과 같이 쓸 수 있다.

$$A^{(n-1)}X = B^{(n-1)} \tag{3.12}$$

여기서 $A^{(n-1)}$은 대각선 상의 원소, 혹은 대각항(*diagonal element*)들 중에 0이 나타나지 않는 상삼각행렬이다. 만일 대각항 중에 0이 있다면 가우스 소거는 실패한 것으로

간주한다. 현재까지 사용한 기본 행작업은 모두 타입 III, 또는 타입 I이다.

이 식의 양변에 $D_{ii} = \frac{1}{a_{ii}^{(n-1)}}$ 인 $n \times n$ 대각행렬을 왼쪽에서 곱하여 얻은 식을 다음과 같이 쓴다. 이것은 타입 II 기본 행작업에 해당한다.

$$A^* X = B^* \tag{3.13}$$

이제 A^*은 대각선 상의 원소가 전부 1인 상삼각행렬이 되었다. 이제 (3.13)의 해는 다음과 같은 *back substitution*을 통하여 구할 수 있다.

$$\begin{aligned} x_n &= b_n^* \\ x_{n-1} &= b_{n-1}^* - a_{n-1,n}^* x_n \\ &\vdots \\ x_1 &= b_1^* - a_{12}^* x_2 - \cdots - a_{1,n}^* x_n \end{aligned} \tag{3.14}$$

가우스 소거란 (3.6)으로부터 (3.14)까지 이르는 과정을 말하는데, 때로는 (3.12)까지만을 말하기도 한다. 혹자는 (3.12)까지의 과정을 *forward elimination*이라고 부르고 가우스 소거는 "forward elimination followed by back substitution"이라고 하기도 한다. 그리고 $A^{(m-1)}$의 대각항들을 1로 만들어 A^*를 얻는 것은 꼭 필요한 것은 아니므로 ((3.14)을 조금만 변형해 주면 x_i들의 값을 계산하는 데 문제가 없으므로) (3.13)은 흔히 생략하기도 한다.

Forward elimination은 적당한 행렬을 A의 왼쪽에 계속 곱함으로써 이루어질 수 있는데 이 점은 back substitution도 마찬가지임을 쉽게 알 수 있을 것이다. 가우스 소거는 A를 이런 방법으로 단위행렬에 이르기까지 변환시키며 B를 $A^{-1}B$로 변환시키는 과정이라고 말할 수 있다. $A^{-1}B$는 방정식 $AX = B$의 해임을 상기하라.

B가 여러 개의 열벡터로 이루어진 행렬일 때도 가우스 소거가 가능하다. 이를 가우스-조던 소거라고 한다. 예를 들어 B가 2개의 열 B^1과 B^2로 이루어진 경우에는 $A^{-1}B = [A^{-1}B^1 | A^{-1}B^2]$를 얻는 것은 2개의 방정식 $AX^1 = B^1$, $AX^2 = B^2$의 해를 동시에 얻는 것으로 볼 수 있다. $B = I_n$에서 시작하면 가우스-조던 소거를 통하여 A^{-1}을 얻게 된다.

가우스-조던 소거 과정에서 A에 일어나는 변화를 정리하면 다음과 같다.

(1) Forward elimination에 의하여 A가 상삼각행렬이 된다. 이때 대각항들은 모두 0 아닌 실수이다. 여기까지는 기본 행작업을 타입 1과 타입 3만 사용해도 된다. (타입 2 기본 행작업을 사용할 수도 있지만 꼭 필요하지는 않다.)

(2) A의 대각항들이 모두 1이 된다. 이 단계에서는 타입 2 기본 행작업만 사용한다.

(3) Back substitution에 의하여 A가 단위행렬이 된다. 이 단계에서는 타입 3 기본 행작업만 사용한다.

<u>정의</u> 3.28 가우스 소거가 불가능한 정방행렬을 특이행렬(*singular matrix*)이라고 한다. 가우스 소거가 가능한 정방행렬은 *nonsingular matrix*이라고 한다. ⊣

<u>정리</u> *3.29* 정방행렬은 *nonsingular*이면이 가역적이다.

(증명). 연습문제 3.27을 사용하면 즉각 알 수 있다. □

<u>연습문제</u> 3.30 다음과 같이 행렬들을 정의한다.

$$A_1 = \begin{bmatrix} -3 & 1 \\ 4 & -2 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -2 & 1 & 3 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}$$

(1) 선형연립방정식 $A_1X = B_1$, $A_1X = C_1$, $A_2X = B_2$, $A_2X = C_2$의 해를 각각 가우스-조던 소거에 의하여 구하여라. (2) 각 $i = 1, 2$에 대하여 B_i과 C_i의 두 열벡터로 이루어진 행렬을 D_i로 두고 $A_iX = D_i$의 해를 가우스-조던 소거에 의하여 구하여라. (3) 가우스-조던 소거에 의하여 계수행렬들의 역행렬을 구한 다음 이것을 B_1, C_1, B_2, C_2에 곱하여 (1), (2)에서 언급한 방정식의 해를 구하여라. ⊣

<u>정리</u> *3.31* (*LDU* 분해) A가 특이행렬이 아닌 정방행렬이면

$$PA = LDU \tag{3.15}$$

를 만족하는 순열행렬 P, 대각항들이 모두 *1*인 하삼각행렬 L, 가역대각행렬 D 및 대각항들이 모두 *1*인 상삼각행렬 U가 존재한다. 게다가 P가 정해진 다음에는 이러한 L, D, U는 유일하게 결정된다. (순열행렬이란 P_{ij}, 즉 타입 I 기본 행작업 행렬들의 곱으로 얻어진 행렬을 뜻한다.)

(증명). 생략 □

3.4 선형사상

정의 3.32 벡터공간 V 에서 벡터공간 W 로 가는 함수 $L : V \to W$ 는 다음의 조건을 만족할 때 선형사상(*linear map*)이라고 한다.

(1) $(\forall c \in \mathbb{R})(\forall \vec{v} \in V)\Big(L(c\vec{v}) = cL(\vec{v})\Big)$

(2) $(\forall \vec{v}_1, \vec{v}_2 \in V)\Big(L(\vec{v}_1 + \vec{v}_2) = L(\vec{v}_1) + L(\vec{v}_2)\Big)$

선형사상은 선형변환(*linear transformation*)이라고도 한다. ⊣

연습문제 3.33 정의 3.32의 조건 (1) and (2)는 다음의 (3.16) 및 (3.17)과 동등함을 보이시오.

$$(\forall c_1, c_2 \in \mathbb{R})(\forall \vec{v}_1, \vec{v}_2)\Big(L(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1L(\vec{v}_1) + c_2L(\vec{v}_2)\Big) \tag{3.16}$$

$$\begin{aligned}&(\forall n \in \mathbb{N})(\forall c_1, \ldots, c_n \in \mathbb{R})(\forall \vec{v}_1, \ldots \vec{v}_n)\\ &\qquad\Big(L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1L(\vec{v}_1) + \cdots + c_nL(\vec{v}_n)\Big)\end{aligned} \tag{3.17}$$

Note: 선형임을 보일 때는 (3.16)을 쓰고, 선형임을 이용할 때는 (3.17)을 사용하는 것이 좋다. ⊣

예제 3.34 선형사상의 가장 간단한 예는 상수함수 $L(\vec{v}) = \vec{0}$ 일 것이다. 공역이 정의역과 일치하는 경우, 혹은 정의역이 공역의 부분공간인 경우에는 항등함수 $L(\vec{v}) = \vec{v}$가 아주 간단한 선형사상의 예가 된다. ⊣

예제 3.35 $A \in \mathbb{R}^{m\times n}$ 일 때 $X \in \mathbb{R}^{n\times 1}$ 을 $AX \in \mathbb{R}^{m\times 1}$ 으로 보내는 함수

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \qquad L_A(X) \stackrel{\text{def}}{=} AX$$

는 선형사상이다. 바로 위의 예에서 소개한 2개의 함수는 이런 형태이다. 모든 선형사상은 (어떤 의미에서) 이러한 형태라는 것을 곧 알게 될 것이다. (See FACT 3.64.) ⊣

연습문제 3.36 L 이 선형사상이면 $L(\vec{0}) = \vec{0}$ 임을 증명하여라. ⊣

연습문제 3.37 $L : V \to W$ 가 선형변환이고, $V' \subseteq V$ 는 V 의 부분공간, $W' \subseteq W$ 는 W 의 부분공간이라 하자. 이때 $L(V')$은 W 의 부분공간이고 $L^{-1}(W')$은 V 의 부분공간임을 보이시오. ⊣

정의 3.38 $L : V \to W$ 가 선형변환일 때 $L^{-1}(\{\vec{0}\})$를 L 의 커늘(*kernel*)이라 하고 $\mathrm{Ker}(L)$로 나타낸다. $\{L(\vec{v}) \mid \vec{v} \in V\}$를 변역(*range*)이라고 하고 $\mathrm{Ran}(L)$로 나타낸다. ⊣

<u>보조정리</u> *3.39* 선형변환 $L : V \to W$ 가 단사라는 것은 $\mathrm{Ker}(L) = \{\vec{0}\}$ 의 필요충분조건이다.

(증명). 나중에 □

<u>연습문제</u> 3.40 선형사상 $L : \mathbb{R}^3 \to \mathbb{R}^3$ 의 커늘의 차원이 0, 1, 2, 3 인 경우에 대한 예를 각각 들고, 그때 $\mathrm{Ker}(L)$ 의 기저와 $\mathrm{Ran}(L)$ 의 기저를 구하여라.

(힌트). 동일한 커늘 U 를 가지는 서로 다른 선형사상 L_1, L_2 가 있을 수 있다. 그리고 이때 $\mathrm{Ran}(L_1) \neq \mathrm{Ran}(L_2)$ 일 수도 있고 아닐 수도 있다. 예를 들어 커늘이 x-축인 경우 (커늘이 1차원임)를 생각해 보자. $L_1((x,y,z)^T) = (0,y,z)^T$, $L_2((x,y,z)^T) = (0,y,y+z)^T$, $L_3((x,y,z)^T) = (y,z,y-z)^T$ 로 두면 $\mathrm{Ker}(L_1) = \mathrm{Ker}(L_2) = \mathrm{Ker}(L_3) = \mathrm{Span}(\vec{\imath})$ 이다. $\mathrm{Ran}(L_1) = \mathrm{Ran}(L_2) = \mathrm{Span}(\vec{\jmath}, \vec{k})$ 이며 $\mathrm{Ran}(L_3) = \mathrm{Span}((1,0,1)^T, (0,1,-1)^T)$ 이다. ⊣

<u>연습문제</u> 3.41 V 는 벡터공간, U 는 V 의 부분집합이라 하자. 선형사상 $L : V \to V$ 에 대한 다음 질문에 답하여라. U 가 선형독립이면 $L(U)$ 는 반드시 선형독립인가? 이것의 역은 어떠한가? ⊣

<u>정리</u> *3.42* $L : V \to W$ 가 선형변환이면

$$\dim(V) = \dim(\mathrm{Ker}(L)) + \dim(\mathrm{Ran}(L))$$

가 성립한다.

(증명). $\mathrm{Ker}(L)$ 의 차원을 ≥ 0 라 하자. 이것의 기저 $\{v_1, \ldots, v_r\}$ 를 취하고 이를 확장하여 V 의 기저 $\{v_1, \ldots, v_r, \ldots, v_n\}$ 를 만든다. 그러면 $\{L(v_{r+1}), \ldots, L(v_n)\}$ 이 $\mathrm{Ran}(L)$ 을 스팬하는 선형독립 집합임이 쉽게 확인된다. □

<u>고찰</u> 3.43 V 의 부분공간 U 에 대한 사영은 정의 2.38에 나타나 있다. 모든 $\vec{p} \in V$ 에 대하여 $\mathrm{Proj}_U \vec{p}$는 U 에 속하는 벡터 중에서 가장 $\vec{p}$와 근사하다는 사실도 이어서 보조정리 2.40에 증명하였다. 다만 그때는 $V = \mathbb{R}^n$ 인 것으로 가정하였는데 증명을 잘 보면 V 가 $\mathbb{R}$ 을 스칼라로 하는 유한차원 내적공간이면 충분하다는 것을 알 수 있을 것이다.

사영 $L(\vec{p}) \overset{\text{def}}{=} \mathrm{Proj}_U \vec{p}$는 선형사상임을 보이겠다. U 의 정규직교기저 하나를 $\{\vec{u}_1, \ldots, \vec{u}_r\}$ 이라 두었을 때

$$L(\vec{p}) = \textstyle\sum_{i=1}^{r} (\vec{u}_i \cdot \vec{p}) \vec{u}_i$$

임을 상기하라. 이제 L 의 선형성의 증명은 다음의 두 부분으로 이루어진다.

(1) 임의의 벡터 $\vec{v} \in V$에 대하여 $\vec{p} \mapsto (\vec{v}\cdot\vec{p})\vec{v}$는 선형사상이다. 이는 내적과 스칼라곱의 기본성질에 의하여 당연히 성립한다.

(2) 두 선형사상의 합은 선형사상이다. 즉 L_1과 L_2가 정의역과 공역을 공유하는 두 선형사상이라면 $(L_1 + L_2)(\vec{p}) \stackrel{\text{def}}{=} L_1(\vec{p}) + L_2(\vec{p})$로 정의되는 함수 $L_1 + L_2$도 선형사상이다. 이 사실은 선형사상의 정의로부터 쉽게 증명된다.

이제 L은 선형사상 r개의 합이므로 선형사상임을 알 수 있다. ⊣

고찰 3.44 여기서 하나 말해 둘 것은 선형사상의 스칼라곱도 선형사상이라는 사실이다. 즉 L이 선형사상이고 c가 스칼라이면 $(cL)(\vec{p}) \stackrel{\text{def}}{=} c(L(\vec{p}))$로 정의되는 함수 cL도 선형사상이며, 이것의 증명은 역시 쉽다.

토의 3.43의 (2)와 바로 위 문단에 의하여 "선형사상들의 집합은 벡터공간을 이룬다."는 사실을 알 수 있을 것이다. 이때 선형사상들은 정의역과 공역을 공유해야 하는 것은 당연하다. 이렇게 이룬 선형사상들의 벡터공간은, 원래 선형사상들의 정의역과 공역이 내적공간이라고 해도 내적공간일 필요는 없다. ⊣

정의 3.45 행렬 $A \in \mathbb{R}^{m\times n}$ 모든 열벡터들로 스팬된 벡터공간을 열공간(*column space*)이라 하고 $\mathrm{ColSpace}(A)$로 나타낸다. A의 모든 행벡터들로 스팬된 벡터공간을 행공간(*row space*)이라 하고 $\mathrm{RowSpace}(A)$로 나타낸다. ⊣

단평 3.46 행렬 $A \in \mathbb{R}^{m\times n}$의 열공간과 행공간은 $L_A(X) = AX,\ (X \in \mathbb{R}^{n\times 1})$로 정의했을 때 각각 다음과 같이 주어진다.

$$\mathrm{ColSpace}(A) = \{AX \mid X \in \mathbb{R}^{n\times 1}\} = \mathrm{Ran}(L_A)$$
$$\mathrm{RowSpace}(A) = \{YA \mid Y \in \mathbb{R}^{1\times m}\}$$
⊣

정리 *3.47* 행렬 $A \in \mathbb{R}^{m\times n}$의 행공간은 $L_A(X) = AX$로 정의했을 때 $\mathrm{Ker}(L_A)$의 직교여공간이다. 즉 $\mathrm{RowSpace}(A) = \mathrm{Ker}(L_A)^\perp$이다. ⊣

이 정리의 증명은 잠시 미루고 먼저 예를 하나 들어 보자.

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

인 경우

$$\mathrm{Ker}(L_A) = \{(x,0,0)^T \mid x \in \mathbb{R}\},$$
$$\mathrm{RowSpace}(A) = \{(0,y,z)^T \mid y, z \in \mathbb{R}\}$$

이 되어 이 둘이 서로 직교여공간임을 알 수 있다. 이제 정리의 증명을 보도록 하자.

(증명). $\mathrm{Ker}(L_A)$의 원소는 $A\vec{x} = \vec{0}$의 해이다. 즉,

$$A\vec{x} = \begin{pmatrix} \vec{A}_1 \cdot \vec{x} \\ \vec{A}_2 \cdot \vec{x} \\ \vdots \\ \vec{A}_m \cdot \vec{x} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.18}$$

따라서 커늘의 원소 $\vec{x}$는 A의 모든 행벡터들과 수직이며 이는 곧 행공간의 직교여공간의 원소임을 의미한다. 역으로 행공간의 직교여공간의 원소를 $\vec{x}$라 하면 이것은 (3.18)을 만족하므로 커늘의 원소가 된다. ⊣

<u>연습문제</u> 3.48 [정리 2.61]에서 유한차원 내적공간 V와 그것의 부분공간 U에 대하여

$$\dim(U) + \dim(U^\perp) = \dim(V)$$

를 보인 바 있는데, 이는 [정리 3.47]과 어떻게 연결되는가? ⊣

<u>정리</u> *3.49* 열공간과 행공간은 차원이 같다. 즉, 모든 행렬 $A \in \mathbb{R}^{m\times n}$에 대해서

$$\dim(\mathrm{ColSpace}(A)) = \dim(\mathrm{RowSpace}(A))$$

이다.

(증명). [정리 3.42]와 [단평 3.46] 및 [정리 3.47]을 사용한다. □

<u>정의</u> 3.50 행렬 A의 열공간의 차원을 A의 랭크(*rank*)라고 하고 $\mathrm{rank}(A)$로 나타낸다. ⊣

<u>예제</u> 3.51 다음과 같은 행렬 $A \in \mathbb{R}^{3\times 4}$가 주어졌을 때

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

$\vec{A}_1 + \vec{A}_2 = \vec{A}_3$이므로 행공간은 $\mathrm{Span}(\vec{A}_1, \vec{A}_2)$이고 이것의 차원은 2가 될 것이다. $\{\vec{A}_1, \vec{A}_2\}$는 선형독립이므로 이것이 곧 행공간의 기저가 되며, 따라서 A의 랭크는 2이다.

커늘의 기저를 찾아 보자. 커늘의 원소는 $\vec{A}_1$과도 수직이고 $\vec{A}_2$와도 수직이다. (그러면 자동으로 $\vec{A}_3$와도 수직이 된다.)

$(1,0,1) \times (0,2,-1) = (-2,1,2)$이므로 $\vec{u}_1 := (-2,1,2,0)^T$는 $\vec{A}_1$, $\vec{A}_2$와 모두 수직이다. 또한 $(1,0,2) \times (0,2,-1) = (-4,1,2)$이므로 $\vec{u}_2 := (-4,1,0,2)^T$는 $\vec{A}_1$, $\vec{A}_2$

와 모두 수직이다. 그러므로 $\vec{u}_1$과 $\vec{u}_2$의 임의의 선형조합은 행공간과 수직이고 따라서 커늘의 원소이다.

이제 역으로 $\vec{A}_1$, $\vec{A}_2$와 모두 수직인 벡터, 즉 커늘의 원소는 모두 $\vec{u}_1$과 $\vec{u}_2$의 선형조합임을 보이겠다. 커늘의 원소를 $\vec{x} := (x, y, z, w)^T$라 했을 때 선형방정식 (3.18)이 만족되어야 하며 이는 다음과 같이 쓸 수 있다.

$$x + z + 2w = 0 \tag{3.19}$$

$$2y - z - w = 0 \tag{3.20}$$

$$x + 2y + w = 0 \tag{3.21}$$

(3.20)으로부터 $w = 2y - z$, (3.21)로부터 $x = -2y - w = -4y + z$를 얻게 되는데 이 두 식을 동시에 만족할 것이 (x, y, z, w)가 이 연립방정식의 해일 필요충분조건임을 쉽게 확인할 수 있다. 그러므로 임의의 커늘의 원소는

$$(-4y + z, y, z, 2y - z)^T = \frac{z}{2}(-2, 1, 2, 0)^T + \left(y - \frac{z}{2}\right)(-4, 1, 0, 2)^T \in \operatorname{Span}(\vec{u}_1, \vec{u}_2)$$

임을 알 수 있다. 따라서 $\{\vec{u}_1, \vec{u}_2\}$는 커늘의 기저이다. 커늘의 차원 2에 A의 랭크 2를 더하면 L_A의 정의역 $\mathbb{R}^4$의 차원이 됨을 확인할 수 있다. ⊣

<u>보조정리</u> *3.52* A, B는 행렬이고 행렬곱 AB가 정의된다고 하자. 그러면

$$\operatorname{rank}(A^T) = \operatorname{rank}(A), \tag{3.22}$$

$$\operatorname{ColSpace}(AB) \subseteq \operatorname{ColSpace}(A), \tag{3.23}$$

$$\operatorname{rank}(AB) \le \min\big(\operatorname{rank}(A), \operatorname{rank}(B)\big) \tag{3.24}$$

이다.

(증명). (3.22)는 $\operatorname{rank}(A^T) = \dim(\operatorname{RowSpace}(A^T)) = \dim(\operatorname{ColSpace}(A)) = \operatorname{rank}(A)$로써 간단히 증명된다.

(3.23)을 증명하기 위하여 $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times r}$로 놓으면

$$\begin{aligned}\operatorname{ColSpace}(AB) &= \{(AB)X \mid X \in \mathbb{R}^{r \times 1}\} \\ &= \{A(BX) \mid X \in \mathbb{R}^{r \times 1}\} \subseteq \{AY \mid Y \in \mathbb{R}^{n \times 1}\} \\ &= \operatorname{ColSpace}(A)\end{aligned}$$

와 같이 쉽게 계산된다.

(3.23)으로부터 $\operatorname{rank}(AB) \le \operatorname{rank}(A)$를 얻을 수 있고, 이어서 (3.22)를 사용하면 $\operatorname{rank}(AB) = \operatorname{rank}\left((AB)^T\right) = \operatorname{rank}\left(B^T A^T\right) \le \operatorname{rank}\left(B^T\right) = \operatorname{rank}(B)$를 얻을 수

있다. ⊣

이제 $A \in \mathbb{R}^{m\times n}$일 때 선형방정식 $A\vec{x} = \vec{b}$의 근의 개수에 대해서 알아 보자. 모든 $\vec{b}$에 대해서 유일한 해를 가진다는 것은 선형사상 $L(\vec{x}) = A\vec{x}$가 전단사인 것과 동등하다. 따라서 $\mathrm{Ker}(L)$의 차원은 0이고 $\mathrm{Ran}(L)$의 차원은 m이어야 한다. 그러면 [정리 3.42]에 의하여 $n = m = \mathrm{rank}(A)$가 되어야 한다.

이러한 경우, 각 $i = 1, \ldots, n$에 대하여 $A\vec{x} = \vec{e_i}$의 유일한 해가 존재할 것인데, 그것을 $\vec{x}^{\,i}$로 나타내면 이들을 열벡터로 하는 행렬 $[\vec{x}^{\,1}, \ldots, \vec{x}^{\,n}]$는 A의 (유일한) 우역행렬(right inverse)이 될 것이다.

이 우역행렬은 좌역행렬이기도 하다는 것, 즉 역행렬이라는 것을 다음과 같이 증명한다. 먼저

$$A \text{ has right inverse } \Leftrightarrow \mathrm{rank}(A) = n \tag{3.25}$$

를 보인다. $\Leftarrow$는 이미 보였다.

$\Rightarrow$를 보이기 위하여는 A가 우역행렬을 가진다고 가정하고 그것을 $A^* \in \mathbb{R}^{n\times n}$로 놓는다. 그러면 $AA^* = I$이므로 $\mathrm{rank}(AA^*) = \mathrm{rank}(I) = n$이 된다. 그러면 이제 (3.24)에 의하여 $\mathrm{rank}(A) = n$이 되어야 한다.

$\Rightarrow$의 증명과정에서 $\mathrm{rank}(A^*) = n$도 알게 되었다.

Theorem 3.53 정방행렬 $A \in \mathbb{R}^{n\times n}$의 우역행렬 A^*는 A의 좌역행렬이기도 하다.

(증명). A의 우역행렬이 존재하므로 (3.25)의 $\Rightarrow$에 의해서 $\mathrm{rank}(A) = n$이다. 우역행렬 A^*의 랭크도 n이므로 A^*는 다시 (3.25)의 $\Leftarrow$에 의해서 그것의 우역행렬 $(A^*)^*$를 가진다.

이제 $A^*A = I$를 다음과 같이 증명한다.

$$(A^*A)\,(A^*A) = A^*\,(AA^*)\,A = A^*IA = A^*A \tag{3.26}$$

$$(A^*A)\,(A^*A)\left(A^*\,(A^*)^*\right) = A^*AI = A^*A \tag{3.27}$$

$$A^*(AA^*)(AA^*)\,(A^*)^* = A^*II\,(A^*)^* = A^*\,(A^*)^* = I \tag{3.28}$$

$$\therefore\ A^*A = I \tag{3.29}$$

(3.27)은 (3.26)의 양변의 오른쪽에 $A^*\,(A^*)^* = I$를 곱하여 얻었다.

(3.28)의 좌변은 (3.27)의 좌변에서 괄호를 고쳐쓴 것이다.

(3.29)의 좌변은 (3.27)의 우변이고 이것은 (3.27)의 좌변과 같고, 이것은 (3.28)의 좌변과 같고, 이것은 다시 (3.28)의 우변 I와 같다.

이것으로써 우역행렬 = 좌역행렬 = 역행렬임을 증명하였다. □

앞서 p42의 [단평 3.9]에서 A와 B가 크기가 같은 정방행렬일 때 $AB = I$이면 $BA = I$라는 사실의 증명은 그리 간단하지 않다고 말하였는데, 이제 [정리 3.53]에서 이를 해결하였다.

역으로 A의 역행렬이 존재하면 모든 $\vec{b}$에 대해서 $A\vec{x} = \vec{b}$가 유일한 해 $\vec{x} = A^{-1}\vec{b}$를 가진다. 현재까지 알아낸 것을 정리하면 다음과 같다.

<u>정리</u> *3.54* 정방행렬 $A \in \mathbb{R}^{n\times n}$이 주어졌을 때 다음은 모두 동등하다.

(*1*) 선형방정식 $A\vec{x} = \vec{b}$가 임의의 $\vec{b}$에 대해서 유일한 해를 가진다.

(*2*) 선형방정식 $A\vec{x} = \vec{b}$가 임의의 $\vec{b}$에 대해서 해를 가진다.

(*3*) 선형방정식 $A\vec{x} = \vec{b}$가 어떤 $\vec{b}$에 대해서 유일한 해를 가진다.

(*4*) 선형방정식 $A\vec{x} = \vec{0}$의 해는 $\vec{0}$ 뿐이다.

(*5*) A의 랭크는 n이다.

(*6*) A의 역행렬이 존재한다. ⊣

<u>연습문제</u> 3.55 A, B, C는 행렬이고 행렬곱 AB와 CA가 정의된다고 하자. 만일 $A \in \mathbb{R}^{n\times n}$가 정방행렬이고 $\mathrm{rank}(A) = n$이면

$$\mathrm{rank}(AB) = \mathrm{rank}(B), \text{ and } \mathrm{rank}(CA) = \mathrm{rank}(C)$$

임을 증명하여라. ⊣

<u>단평</u> 3.56 위의 [연습문제 3.55]에 의하면 행렬에 기본행작업을 실행하면 랭크가 바뀌지 않는다. 왜냐하면 기본행작업의 실행은 기본행작업 행렬을 왼쪽에서 곱하는 것과 같고, 기본행작업 행렬은 가역행렬이기 때문이다.

사실 기본행작업이 랭크를 보존한다는 것은 구태여 위에서 말한 "행렬곱의 랭크" 개념을 사용하지 않고도 설명할 수 있다. 행렬의 행공간은 기본행작업에 의하여 바뀌지 않기 때문이다. $3 \times m$ 행렬을 예로 들어 설명하자면

$$\begin{aligned} \mathrm{Span}(A_1, A_2, A_3) &= \mathrm{Span}(A_2, A_1, A_3) \\ &= \mathrm{Span}(A_1, A_2, A_3 + kA_1) \\ &= \mathrm{Span}(kA_1, A_2, A_3), \quad (k \neq 0) \end{aligned}$$

등이 성립하기 때문이다.

이상의 논의는 기본열작업에 대해서도 마찬가지로 성립한다. ⊣

<u>정의</u> 3.57 사다리꼴 행렬(*echelon form matrix*)[†]은 다음의 조건을 만족하는 행렬을 뜻한다.

(1) 각 행의 피봇(*pivot*) 원소(0 아닌 원소 중 가장 왼쪽에 나타나는 것)는 그 위쪽에 있는 행의 피봇 원소보다 오른쪽에 있다.

(2) 피봇 원소가 없는 행, 즉 0-벡터인 행들은 행렬의 아래쪽에 모여 있다.

기약 사다리꼴 행렬(*reduced row echelon form matrix*)는 사다리꼴 행렬로서 다음의 두 조건을 추가로 만족하는 것이다.

(3) 모든 피봇 원소는 1이다.

(4) 피봇 원소를 포함한 열의 모든 원소는 그 피봇을 포함한 행을 제외한 모든 행에서 0이다.

아래에 사다리꼴 및 기약 사다리꼴 행렬의 예를 들어 보았다. ⊣

$$\begin{bmatrix} 1 & 3 & 0 & 2 & -1 \\ 0 & 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 3 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

기약 사다리꼴 행렬의 대표적인 예로 단위행렬이 있다. 사다리꼴 행렬의 랭크는 피봇 원소를 가진(즉, $\vec{0}$아닌) 행들의 개수와 일치함을 쉽게 알 수 있다.

<u>고찰</u> 3.58 $A \in \mathbb{R}^{n\times n}$, $\vec{b} \in \mathbb{R}^n$일 때 방정식 $A\vec{x} = \vec{b}$의 해의 개수는 $\mathrm{rank}(A) = n$일 때는 정확히 1개임을 보였다. 여기서는 $\mathrm{rank}(A) < n$인 경우에 해의 숫자에 대해서 알아 보자. 결과만 말하자면 해의 개수는 0이거나 무한대이다. 예를 들어 딱 2개의 해를 가지는 경우는 없다.

아래의 행렬 A, $\vec{b}_1$, $\vec{b}_2$에 대하여 방정식 $A\vec{x} = \vec{b}$는 $\vec{b} = \vec{b}_1$일 때는 해를 하나도 가지지 못하고, $\vec{b} = \vec{b}_2$일 때는 해를 무한히 많이 가짐을 알 수 있을 것이다.

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, \quad \vec{b}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

이 예에서 $\mathrm{rank}(A) = 1 < n$인데 이는 [정리 3.54]에 의하여 당연히 예상되었던 것이다.

이제 $\mathrm{rank}(A) < n$를 가정하고 $A\vec{x} = \vec{b}$의 해의 개수가 0 또는 무한대임을 증명해 보자. $\vec{b} = \vec{0}$인 경우는 쉽다. $A\vec{x} = \vec{0}$의 해들의 집합은 곧 $\mathrm{Ker}(L_A)$이고 이것의 차원은

[†]An <u>echelon</u> is a military formation in which soldiers, vehicles, ships, or aircraft follow each other but are spaced out sideways so that they can see ahead.

$n - \mathrm{rank}(A) > 0$이므로 해가 무한히 많다. 일반적인 경우 $\vec{b} \neq \vec{0}$를 생각해 보자. 두 가지 경우가 있다. $A\vec{x} = \vec{b}$의 해가 존재할 경우와 그렇지 않을 경우다. 우리가 보여야 할 것은 전자의 경우 해가 무한히 많이 존재한다는 것이다.

해가 하나라도 존재한다고 가정하고 그것을 $\vec{x}_0$라 둔다. 그리고 $A\vec{x} = \vec{0}$의 무한히 많은 해 중에서 임의로 하나를 택하여 $\vec{y}$로 두면 $\vec{x}_0 + \vec{y}$는 $A\vec{x} = \vec{b}$의 해가 됨이 쉽게 확인된다: $A(\vec{x}_0 + \vec{y}) = A\vec{x}_0 + A\vec{y} = \vec{b} + \vec{0} = \vec{b}$. □

3.5 좌표벡터와 표상행렬

이 절에서는 벡터를 나타낼 때 위의 화살표를 사용하지 않고 그냥 소문자 u, v, w 등을 사용할 것이다. 다만 이들이 $\mathbb{R}^n$의 원소인 경우에는 화살표를 사용하며 행렬곱에서 사용하기 쉽도록 모두 열벡터인 것으로 간주한다. 예를 들어 $\vec{e}_1 = (1,0)^T$, $\vec{e}_2 = (0,1)^T$이다.

정의 3.59 벡터공간의 기저에 순서를 준 것을 순서기저(*ordered basis*)라고 한다. V의 순서기저 $B \overset{\text{def}}{=} (v_1, \ldots, v_n)$가 주어졌다 하자. 각 벡터 $v \in V$에 대해서 이를 $v = a_1v_1 + \cdots + a_nv_n$과 같은 선형조합으로 나타냈을 때 열벡터 $(a_1, \ldots, a_n)^T \in \mathbb{R}^n$를 "$v$의 순서기저 B에 대한 좌표벡터(*coordinate vector*)"라고 하고 $M_B(v)$로 나타낸다. ⊣

예제 3.60 $B_1 := (\vec{e}_1, \vec{e}_2)$, $B_2 := (\vec{e}_2, \vec{e}_1)$, $v = (3,4)^T$로 두면 $M_{B_1}(v) = (3,4)^T$이고 $M_{B_2}(v) = (4,3)$이다. ⊣

연습문제 3.61 $\mathbb{R}^2$의 두 순서기저

$$B_1 = ((1,1)^T, (-1,1)^T) \overset{\text{def}}{=} \{v_1, v_2\},$$
$$B_2 = ((1,2)^T, (2,1)^T) \overset{\text{def}}{=} \{w_1, w_2\}$$

가 주어졌을 때 다음 물음에 답하여라. 단 $v = (3,4)^T$이다.

(1) $M_{B_1}(v)$과 $M_{B_2}(v)$를 구하여라. (힌트). B_1의 원소들을 열벡터로 가지는 행렬을 A_v, B_2의 원소들을 열벡터로 가지는 행렬을 A_w라고 했을 때 $v = A_vM_{B_1}(v) = A_wM_{B_2}v$.

(2) $M_{B_2}(v) = AM_{B_1}(v)$을 만족하는 $A \in \mathbb{R}^{2\times 2}$를 구하여라. (힌트). A는 A_v와 A_w를 이용해서 나타낼 수 있다.

(3) 다른 v에 대해서도 (2)에서 구한 행렬 A를 사용할 수 있는가? ⊣

연습문제 3.62 위의 연습문제에서 정의했던 벡터 v_1, v_2, w_1, w_2에 대해서 다음의 식이 성립하는 $(a_{ij})_{ij}$, $(i, j = 1, 2)$를 구하여라.

$$v_1 = a_{11}w_1 + a_{21}w_2,$$
$$v_2 = a_{12}w_1 + a_{22}w_2.$$

$A := (a_{ij})_{ij}$로 두면 모든 $v \in \mathbb{R}^2$에 대해서 $M_{B_2}(v) = AM_{B_1}(v)$가 성립함을 보여라. (즉 이 A는 [연습문제 3.61].(2)에서 구했던 행렬 A와 같다.) 이러한 A를 $M_{B_2}^{B_1}$, 또는 $M_{B_2}^{B_1}(id)$로 나타낸다. ⊣

<u>정의</u> 3.63 $L : V \to W$이 선형사상이고 $B_1 \stackrel{\text{def}}{=} (v_1, \ldots, v_n)$은 V의 순서기저, $B_2 \stackrel{\text{def}}{=} (w_1, \ldots, w_m)$는 W의 순서기저라 하자. 각 $L(v_i)$를 다음과 같이 B_2의 선형조합으로 나타내었을 때

$$L(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots a_{m1}w_m,$$
$$L(v_2) = a_{12}w_1 + a_{22}w_2 + \cdots a_{m2}w_m,$$
$$\vdots$$
$$L(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots a_{mn}w_m.$$

행렬 $M_{B_2}^{B_1}(L) \stackrel{\text{def}}{=} A \in \mathbb{R}^{m\times n}$을 $(A)_{ij} = a_{ij}$로 정의한다. $M_{B_2}^{B_1}(L)$는 (B_1, B_2) 기저환경에서 L의 표상행렬(*representation matrix*)이다. ⊣

Fact 3.64 $L : V \to W$이 선형사상이고 $B_1 \stackrel{\text{def}}{=} (v_1, \ldots, v_n)$은 V의 순서기저, $B_2 \stackrel{\text{def}}{=} (w_1, \ldots, w_m)$은 W의 순서기저라 했을 때 모든 $v \in V$에 대하여 다음의 등식이 성립한다.

$$M_{B_2}(L(v)) = M_{B_2}^{B_1}(L)M_{B_1}(v) \tag{3.30}$$

(3.30)을 "표상행렬과 좌표벡터들의 관계식"이라고 부른다.

$L = id_V$, 즉 항등함수 $v \mapsto v$일 때는 (3.30)은

$$M_{B_2}(v) = M_{B_2}^{B_1}(id_V)M_{B_1}(v) \tag{3.31}$$

가 된다. 이 식은 좌표계의 변환에 따른 좌표벡터의 변화를 말해준다.

(증명). (3.30)을 증명하기 위하여 $v = c_1v_1 + \cdots + c_nv_n \in V$가 주어졌다 하자. 그러면 좌표벡터의 정의에 의해서 $M_{B_1}(v) = (c_1, \ldots, c_n)^T$이다. $L(v) = c_1L(v_1) + \cdots + c_nL(v_n)$

이므로 $M_{B_2}^{B_1}(L) \stackrel{\text{def}}{=} A$ 의 정의에 의하여

$$\begin{aligned} L(v) &= c_1(a_{11}w_1 + a_{21}w_2 + \cdots a_{m1}w_m) + \\ & \quad c_2(a_{12}w_1 + a_{22}w_2 + \cdots a_{m2}w_m) + \\ & \quad \vdots \\ & \quad c_n(a_{1n}w_1 + a_{2n}w_2 + \cdots a_{mn}w_m) \\ &= \textstyle\sum_{i=1}^{m}(c_1a_{i1} + \cdots + c_na_{in})w_i \end{aligned}$$

가 된다. 여기서 기저 원소 w_i 의 계수를 보면

$$c_1a_{i1} + \cdots + c_na_{in} \tag{3.32}$$

이고 이것이 $M_{B_2}(L(v))$ 의 (정의에 의하여) i 번째 원소가 된다.

한편 $M_{B_2}^{B_1}(L)M_{B_1}(v) = A\,(c_1, \ldots, c_n)^T$ 의 i 번째 원소 $A_i \cdot \vec{c}$는 (3.32)와 일치한다. 이것으로써 (3.30)이 성립함이 확인되었다. □

<u>예제</u> 3.65 $\mathbb{R}^2$ 위의 선형사상 $L\left((x,y)^T\right) = (x+y, -x+2y)^T$ 의 표상행렬 $M_{B_2}^{B_1}(L)$ 을 얻어 보자. 기저환경은 다음과 같이 4가지 경우를 각각 사용한다.

(i) $B_1 = B_2 = (\vec{e}_1, \vec{e}_2)$ (표준순서기저)

(ii) $B_1 = (\vec{e}_1, \vec{e}_2)$, $B_2 = ((1,1)^T, (-1,1)^T) \stackrel{\text{def}}{=} (v_1, v_2)$

(iii) $B_1 = (v_1, v_2)$, $B_2 = (\vec{e}_1, \vec{e}_2)$

(iv) $B_1 = (v_1, v_2)$, $B_2 = ((1,2)^T, (2,1)^T) \stackrel{\text{def}}{=} (w_1, w_2)$

<u>경우 (i)</u>: $L(\vec{e}_1) = L((1,0)^T) = (1,-1)^T = 1\vec{e}_1 - 1\vec{e}_2 = a_{11}\vec{e}_1 + a_{21}\vec{e}_2$, $L(\vec{e}_2) = (1,2)^T = 1\vec{e}_1 + 2\vec{e}_2 = a_{12}\vec{e}_1 + a_{22}\vec{e}_2$.

$$\therefore\ M_{B_2}^{B_1}(L) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

(3.30)은 다음과 같이 확인된다. $v = (x,y)^T = x\vec{e}_1 + y\vec{e}_2$ 일 때

$$M_{B_1}(v) = (x,y)^T$$

$$L(v) = (x+y, -x+2y)^T = (x+y)\vec{e}_1 + (-x+2y)\vec{e}_2$$

$$M_{B_2}(L(v)) = (x+y, -x+2y)^T$$

$$\therefore\ M_{B_2}^{B_1}(L)M_{B_1}(v) = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ -x+2y \end{bmatrix} = M_{B_2}(L(v))$$

<u>경우 (ii)</u>: $L(\vec{e}_1) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = a_{11}v_1 + a_{21}v_2 = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 1\begin{bmatrix} -1 \\ 1 \end{bmatrix}$
$L(\vec{e}_2) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = a_{12}v_1 + a_{22}v_2 = \frac{3}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$\therefore\ M_{B_2}^{B_1}(L) = \begin{bmatrix} 0 & \frac{3}{2} \\ -1 & \frac{1}{2} \end{bmatrix}$$

(3.30)은 다음과 같이 확인된다. $v = (x, y)^T = x\vec{e}_1 + y\vec{e}_2$ 일 때

$$M_{B_1}(v) = (x, y)^T$$
$$L(v) = \begin{bmatrix} x+y \\ -x+2y \end{bmatrix} = \frac{3y}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-x+\tfrac{y}{2})\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$
$$M_{B_2}(L(v)) = (\tfrac{3y}{2}, -x+\tfrac{y}{2})^T$$
$$\therefore\ M_{B_2}^{B_1}(L)M_{B_1}(v) = \begin{bmatrix} 0 & \frac{3}{2} \\ -1 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{3y}{2} \\ -x+\frac{y}{2} \end{bmatrix} = M_{B_2}(L(v))$$

<u>경우 (iii)</u>: $L(\vec{v}_1) = L\left((1,1)^T\right) = (2,1)^T = a_{11}\vec{e}_1 + a_{21}\vec{e}_2 = 2\vec{e}_1 + 1\vec{e}_2$
$L(\vec{v}_2) = L\left((-1,1)^T\right) = (0,3)^T = a_{12}\vec{e}_1 + a_{22}\vec{e}_2 = 0\vec{e}_1 + 3\vec{e}_2$

$$\therefore\ M_{B_2}^{B_1}(L) = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$$

(3.30)은 다음과 같이 확인된다. $v = (x, y)^T = \frac{x+y}{2}\vec{v}_1 + \frac{-x+y}{2}\vec{v}_2$ 일 때

$$M_{B_1}(v) = (\tfrac{x+y}{2}, \tfrac{-x+y}{2})^T$$
$$L(v) = (x+y, -x+2y)^T = (x+y)\vec{e}_1 + (-x+2y)\vec{e}_2$$
$$M_{B_2}(L(v)) = (x+y, -x+2y)^T$$
$$\therefore\ M_{B_2}^{B_1}(L)M_{B_1}(v) = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} \frac{x+y}{2} \\ \frac{-x+y}{2} \end{bmatrix} = \begin{bmatrix} x+y \\ -x+2y \end{bmatrix} = M_{B_2}(L(v))$$

<u>경우 (iv)</u>: $L(\vec{v}_1) = L\left((1,1)^T\right) = (2,1)^T = a_{11}w_1 + a_{21}w_2 = 0(1,2)^T + 1(2,1)^T$
$L(\vec{v}_2) = L\left((-1,1)^T\right) = (0,3)^T = a_{12}w_1 + a_{22}w_2 = 2(1,2)^T - 1(2,1)^T$

$$\therefore\ M_{B_2}^{B_1}(L) = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}$$

(3.30)은 다음과 같이 확인된다. $v = (x, y)^T = \frac{x+y}{2}v_1 + \frac{-x+y}{2}v_2$ 일 때

$$
\begin{aligned}
&M_{B_1}(v) = (\tfrac{x+y}{2}, \tfrac{-x+y}{2})^T \\
&L(v) = \begin{bmatrix} x+y \\ -x+2y \end{bmatrix} = (-x+y)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + x\begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-x+y)w_1 + xw_2 \\
&M_{B_2}(L(v)) = (-x+y, x)^T \\
&\therefore\ M_{B_2}^{B_1}(L)M_{B_1}(v) = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \frac{x+y}{2} \\ \frac{-x+y}{2} \end{bmatrix} = \begin{bmatrix} -x+y \\ x \end{bmatrix} = M_{B_2}(L(v))
\end{aligned}
$$

⊣

고찰 3.66 [예제 3.65].(iv) 에서 표상행렬 $M_{B_2}^{B_1}(L)$을 얻는 체계적인 방법을 생각해 보자. $L((x,y)^T) = (x+y, -x+2y)^T$ 는 행렬 A_L 을 도입하여 다음과 같이 쓸 수 있다.

$$
L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ -x+2y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = A_L\begin{pmatrix} x \\ y \end{pmatrix}
$$

$M_{B_2}^{B_1}(L)$을 얻기 위한 식

$$
\begin{aligned}
&L(v_1) = a_{11}w_1 + a_{21}w_2, \\
&L(v_2) = a_{12}w_1 + a_{22}w_2.
\end{aligned}
$$

는 행렬 $A_w := (w_1 \mid w_2)$를 도입하여 다음과 같이 쓸 수 있다.

$$
\begin{aligned}
&A_L\begin{pmatrix} 1 \\ 1 \end{pmatrix} = a_{11}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_{21}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = A_w\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \\
&A_L\begin{pmatrix} -1 \\ 1 \end{pmatrix} = a_{12}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_{22}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = A_L\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}. \\
&\therefore\ A_L\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = A_w\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A_w M_{B_2}^{B_1}(L)
\end{aligned}
$$

여기서 다시 행렬 $A_v := (v_1 \mid v_2)$을 도입하면 위의 식은

$$
A_L A_v = A_w M_{B_2}^{B_1}(L)
$$

이 되고 따라서

$$
M_{B_2}^{B_1}(L) = A_w^{-1} A_L A_v \tag{3.33}
$$

를 얻을 수 있다. (3.33)에 의하여 표상행렬을 계산하면 [예제 3.65].(iv) 에서 얻었던

결과와 일치함을 지오지브라 등의 소프트웨어를 이용하여 쉽게 확인할 수 있을 것이다.

(3.33)은 표상행렬을 구할 때 편리하게 사용할 수 있는 공식이다. 단, 이 공식은 정의역과 공역이 $\mathbb{R}^n$ 또는 이것의 부분공간일 때만 적용할 수 있다. 왜냐하면 A_v, A_w 등은 순서기저의 원소들을 열벡터로 가지는 행렬이며 이때 열벡터는 $\mathbb{R}^n$의 원소이기 때문이다. ⊣

<u>연습문제</u> 3.67 이 문제에서는 [예제 3.65]에서 사용했던 벡터 v_1, v_2, w_1, w_2와 선형사상 L을 그대로 사용한다. 계산에 지오지브라를 활용하기를 권한다.

(1) [예제 3.65]에 경우 (v)를 추가한다: $B_1 = (v_1, v_2)$, $B_2 = (w_1, w_2)$. 이 경우에 대해서 행렬 $M_{B_1}^{B_2}(L)$을 구하고

$$(\forall v \in \mathbb{R}^2)\Big(M_{B_1}(L(v)) = M_{B_1}^{B_2}(L)M_{B_2}(v)\Big)$$

가 됨을, 즉 $v = (x, y)^T$로 놓고 $M_{B_2}(v)$와 $M_{B_1}(L(v))$를 구하여 위의 등식이 성립함을 확인하여라.

(2) $L : \mathbb{R}^2 \to \mathbb{R}^2$를 $L^2(v) := L(L(v))$로 정의한다. B_3를 표준기저라고 했을 때 $M_{B_3}^{B_3}(L)$과 $M_{B_3}^{B_3}(L^2)$을 구하여라. 그리고 $M_{B_1}^{B_1}(L)$, $M_{B_1}^{B_1}(L^2)$, $M_{B_2}^{B_2}(L)$, $M_{B_2}^{B_2}(L^2)$을 구하여라.

(3) $L_1 = L$로 두고 $L_2\big((x, y)^T\big) = (x - 2y, 2x + y)^T$, $(L_1 \circ L_2)(v) = L_1(L_2(v))$로 정의한다. 이때 $M_{B_3}^{B_1}(L_1 \circ L_2)$, $M_{B_2}^{B_1}(L_2)$, $M_{B_3}^{B_2}(L_1)$을 각각 구하여라.

(4) $M_{B_2}^{B_1}(id)$, $M_{B_1}^{B_2}(id)$, $M_{B_2}^{B_1}(L)$, $M_{B_1}^{B_2}(L)$, $M_{B_3}^{B_3}(L^{-1})$, $M_{B_1}^{B_2}(L^{-1})$, $M_{B_2}^{B_1}(L^{-1})$을 각각 구하여라. ⊣

<u>정리</u> *3.68* 선형사상의 합성은 표상행렬의 곱에 대응되고 선형사상의 역함수는 표상행렬의 역행렬에 대응된다. 즉, 모든 선형사상 L_1, L_2 L에 대하여

$$M_{B_3}^{B_1}(L_1 \circ L_2) = M_{B_3}^{B_2}(L_1)M_{B_2}^{B_1}(L_2), \tag{3.34}$$

$$\Big(M_{B_2}^{B_1}(L)\Big)^{-1} = M_{B_1}^{B_2}(L^{-1}). \tag{3.35}$$

가 성립한다. 단, $L_1 \circ L_2$와 L^{-1}이 존재한다고 가정한다.

(증명). (3.34)는 표상행렬과 좌표벡터의 정의 및 행렬곱의 결합법칙에 의하여 성립한다. (3.35)는 $L_1 = L$, $L_2 = L^{-1}$, $B_3 = B_1$를 (3.34)에 적용하면, $M_{B_1}^{B_1}(id) = I$를 이용하여 얻을 수 있다. □

<u>따름정리</u> *3.69* $L : V \to V$가 선형사상이고 B_1, B_2가 V의 순서기저이면 어떤 가역행렬 N에 대해서

$$M_{B_1}^{B_1}(L) = N^{-1} M_{B_2}^{B_2}(L)\, N$$

가 성립한다.

(증명). $N = M_{B_2}^{B_1}(id_V)$로 두면 된다. □

<u>연습문제</u> 3.70 $\mathbb{R}^n$의 표준순서기저를 B_1이라 하고 $B_2 \stackrel{\text{def}}{=} (v_1, \ldots, v_n)$을 $\mathbb{R}^n$의 어떤 순서기저라 하면 v_i를 i-th column vector로 가지는 정방행렬은 $M_{B_1}^{B_2}(id_V)$임을 보이시오. ⊣

<u>단평</u> 3.71 좌표벡터와 표상행렬은 선형대수에서 가장 중요한 부분 중 하나이며 우리가 행렬을 공부해야 하는 이유를 말해 준다. 즉, 선형사상은 (기저가 정해지면) 행렬과 1대 1 대응이 된다는 것이다. 선형사상의 합성은 행렬곱에 대응되고, 선형사상의 역함수는 역행렬에 대응된다.

선형사상은 추상적 관념이고 행렬은 구체적 실체이다. 추상의 세계, 즉 관념은 그것의 표상이 있어야 우리가 다룰 수 있다. 특히 남과 의견을 나누고 알아낸 사실을 기록할 때 표상이 없이는 아무것도 할 수 없을 것이다. 행렬은 선형사상의 훌륭한 표상이다. ⊣

<u>보조정리</u> *3.72* $\{v_1, \ldots, v_n\}$이 어떤 벡터공간의 기저라 하자. $n \times n$ 행렬 $A \stackrel{\text{def}}{=} (a_{ij})_{ij}$에 대하여

$$\begin{aligned}
w_1 &= a_{11}v_1 + a_{21}v_2 + \cdots a_{n1}v_n,\\
w_2 &= a_{12}v_1 + a_{22}v_2 + \cdots a_{n2}v_n,\\
&\vdots\\
w_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots a_{nn}v_n.
\end{aligned} \tag{3.36}$$

로 놓으면 $\{w_1, \ldots, w_n\}$은 A가 가역행렬일 때, 그리고 이때만 기저를 이룬다.

(증명). ($\Leftarrow$) $B_1 \stackrel{\text{def}}{=} \{v_1, \ldots, v_n\}$, $B_2 \stackrel{\text{def}}{=} \{w_1, \ldots, w_n\}$로 놓고 B_2가 기저라고 가정하자. 그러면

$$M_{B_2}^{B_1}(I) M_{B_1}^{B_2}(I) = M_{B_2}^{B_2}(II) = M_{B_2}^{B_2}(I) = I = M_{B_1}^{B_1}(I) = M_{B_1}^{B_2}(I) M_{B_2}^{B_1}(I)$$

이므로 $A = M_{B_1}^{B_2}(I)$는 가역행렬이며, 이것의 역행렬은 $M_{B_2}^{B_1}(I)$임을 알 수 있다.

(⇒) 역으로 A가 가역행렬이라고 가정하자. 임의의 벡터

$$v = a_1 v_1 + \cdots + a_n v_n \tag{3.37}$$

이 주어졌을 때

$$v = b_1 w_1 + \cdots + b_n w_n \tag{3.38}$$

을 만족하는 b_i들이 유일하게 존재함을 보이면 된다. (3.36)을 (3.38)에 대입한 다음 (3.37)와 비교하면 $\vec{a} = A\vec{b}$를 얻는다. 이 식은 $\vec{b} = A^{-1}\vec{a}$와 동등하다. □

3.6 디터미넌트

<u>정의</u> 3.73 각 $n \geq 1$에 대해서 함수

$$\det : \mathbb{R}^{n\times n} \to \mathbb{R}$$

를 아래의 성질을 만족하는 함수로 정의한다.

(1) det는 행렬의 첫 행에 대한 선형사상이다. 즉

① $A, B \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}$이고

$$B_1 = cA_1 \text{ and } (\forall i = 2, \ldots, n)(A_i = B_i)$$

이면 $\det(B) = c\det(A)$이다.

② $A, B, C \in \mathbb{R}^{n\times n}$이고

$$C_1 = A_1 + B_1 \text{ and } (\forall i = 2, \ldots, n)(A_i = B_i = C_i)$$

이면 $\det(C) = \det(A) + \det(B)$이다.

(2) $A \in \mathbb{R}^{n\times n}$에서 임의의 두 행을 취한 뒤 서로 바꾸어 $B \in \mathbb{R}^{n\times n}$를 얻으면

$$\det(B) = -\det(A)$$

이다.

(3) $\det(I_n) = 1$.

A가 행렬일 때 $\det(A)$를 A의 디터미넌트(*determinant*)라고 한다. $\det(A)$ 대신 $|A|$로 나타내기도 한다. ⊣

이 정의가 유용하려면 다음의 2가지가 준비되어야 한다. (1) $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$는 잘 정의된(well-defined) 함수임을 증명한다. (2) 주어진 행렬의 디터미넌트를 계산하는 알고리듬을 찾는다.

이 둘은 다음의 정리에서 해결된다.

정리 *3.74* 행렬 $A \in \mathbb{R}^{n\times n}$의 디터미넌트는 아래의 성질들을 가진다.

(*4*) 동일한 두 개의 행이 존재하는 행렬의 디터미넌트는 *0*이다. 어떤 행이 다른 행의 스칼라곱인 행렬의 디터미넌트는 *0*이다.

(*5*) 하나의 행의 스칼라곱을 다른 행에 더하거나 빼어 새로운 행렬을 얻으면 디터미넌트는 변하지 않는다.

(*6*) 디터미넌트는 각 행에 대한 선형사상이다.

(*7*) 대각행렬의 디터미넌트는 대각항들의 곱이다. 삼각행렬에서도 디터미넌트는 대각항들의 곱이다. (이제 가우스 소거를 이용하면 디터미넌트의 존재성과 유일성 및 알고리듬이 얻어진다.)

(*8*) $\det A \neq 0$이면이 A는 가역적이면이 $\operatorname{rank}(A) = n$이다.

(*9*) $\det AB = \det A \det B$.

(*10*) $\det A^T = \det A$ (*LDU*-분해를 이용한다.)

(증명). 증명의 쉬운 부분은 연습문제로 독자에게 넘긴다.

(7)은 $n = 3$인 경우에 대해서만 아래에 증명을 보였다. $a \neq 0, b \neq 0, c \neq 0$인 대각행렬의 경우는 다음과 같이 증명된다.

$$\begin{vmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{vmatrix} = a \begin{vmatrix} 1 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{vmatrix} = ab \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{vmatrix} = abc \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = abc$$

삼각행렬의 경우에는 타입 3 기본행 작업에 의하여 대각행렬로 변환할 수 있으므로 (2)에 의하여 증명된다.

(8)은 [정리 3.29]와 [정리 3.54]를 사용하여 증명할 수 있다. 전자는 가우스-조던 소거와 가역성에 대한 정리이고 후자는 랭크와 가역성에 대한 정리이다.

(9)는 $\det(B) = 0$이면 $\operatorname{rank}(AB) \leq \operatorname{rank}(B) < n$이므로 $\det AB = 0$이다. $\det B \neq 0$이면 $f(A) \stackrel{\text{def}}{=} \det AB / \det B$는 디터미넌트 정의의 3 성질을 모두 가진다.

이 사실의 증명에는 [단평 3.3.(3)]에 서술된 행렬곱을 보는 3번째 관점, 즉 "$AB = C$ 일 때 행백터 C_i는 행백터 A_i와 B의 곱이며 A의 다른 행들과는 하등 상관이 없다. C_i는 B의 행벡터들의 선형조합이며 A_i의 항들을 계수로 가진다."이 이용된다. 따라서 $f(A) = \det A$가 된다.

(10)은 [정리 3.31]을 사용하면 곧 알 수 있다. 삼각정방행렬 A에 대하여는 $\det(A^T) = \det(A)$라는 사실을 이용한다. □

고찰 3.75 이 부분은 설명을 간략하게 처리하였으므로 시간이 부족한 학생은 유도 과정을 완벽하게 이해하지 않더라도 cofactor expansion을 써서 디터미넌트를 계산하는 방법, 그리고 Cramer's rule만 익히면 되리라 본다.

$n \times n$-행렬의 디터미넌트는 각 행에 대한 선형사상이라는 사실을 이용하여 각 행에 0 아닌 항이 기껏해야 한 개씩만 있는 n^n개의 행렬의 각각의 디터미넌트의 합으로 나타낼 수 있다. 예를 들어 $n = 3$인 경우

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + \begin{vmatrix} 0 & 2 & 0 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

이며, 우변에 있는 3개의 디터미넌트는 각각 2번째 행에 대하여 전개하여 3개의 디터미넌트의 합이 된다. 현재까지 얻은 9개의 디터미넌트의 각각은 다시 3번째 행에 대하여 전개하여 3개의 디터미넌트의 합으로 표시된다. 그리하여 원래의 디터미넌트는 $3^3 = 27$개의 디터미넌트의 합이며 이 디터미넌트들은 각 행에 0 아닌 항이 1개 이하가 있는 "원자" 행렬의 디터미넌트이다.

그런데 n^n개의 디터미넌트 중에 많은 것들의 값이 0이다. 어떤 특정한 열의 모든 항이 0인 행렬의 디터미넌트는 0이기 때문이다. 조금만 생각해 보면 n^n개의 행렬 중에서 $\{1, \ldots, n\}$의 "순열에 대응하는" $n!$개만 남기고는 모두 0임을 알 수 있다. 순열 $\sigma \in S_n$에 대응하는 원자 행렬은 i번째 행의 $a_{i,\sigma(i)}$만 그대로고 나머지는 모두 0인 행렬이다. 그리고 이 행렬들은 행을 여러 번 맞교환 하면 대각행렬로 변환되며, 두 행의 맞교환은 디터미넌트의 부호만을 바꾸므로 우리는 결국 다음의 식을 얻을 수 있다.

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}. \tag{3.39}$$

여기서 $\operatorname{sign}(\sigma)$는 σ가 우순열일 때는 1, 기순열일 때는 -1이다. 디터미넌트의 정의로부터 (3.39)을 유도하였는데, 역으로 (3.39)과 같이 정의된 det는 디터미넌트 정의의 3가지 성질를 만족함을 어렵지 않게 확인할 수 있다.

이 확인과정에서 *cofactor expansion*의 개념이 자연스럽게 도출된다. Cofactor expansion을 사용하면 *Cramer's rule*이 쉽게 유도된다. ⊣

각 $1 \le i, j \le n$에 대하여 A의 (i,j)-*minor*는 A에서 i째 행과 j째 열을 삭제해서 얻은 $(n-1) \times (n-1)$ 행렬을 뜻하며 $A_{(ij)}$로 나타낸다.

A의 여인수 행렬(*cofactor matrix*) $A_{\mathrm{cof}} \in \mathbb{R}^{n \times n}$는 ij-항이 $(-1)^{i+j} \det A_{(ij)}$인 행렬로 정의된다. 이때 $\det A_{(ij)}$의 부호 $(-1)^{i+j}$는 다음과 같은 *checkerboard pattern*으로 시각화 하면 이해하기 쉽다. $n = 4$인 경우에 대한 그림인데 다른 값의 n에 대해서도 쉽게 짐작할 수 있을 것이다.

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

여인수 전개(*Cofactor expansion*)란 $\det A$를 다음과 같은 방법으로 계산하는 것을 뜻한다. A_{cof}의 ij-항을 c_{ij}로 나타내기로 했을 때, i째 행을 따른 여인수 전개는

$$\det A = \sum_{j=1}^{n} a_{ij} c_{ij}$$

i의 값으로 $1, \ldots, n$ 중 어느 것을 택하더라도 결과는 같다. j째 행을 따른 여인수 전개는

$$\det A = \sum_{i=1}^{n} a_{ij} c_{ij}$$

j의 값으로 $1, \ldots, n$ 중 어느 것을 택하더라도 결과는 같다.

아래에 여인수 전개의 예를 보였다. 먼저 첫번째 행을 따른 계산을 보였다.

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 1 \cdot (-3) - 2 \cdot (-6) + 3 \cdot (-3) = -3 + 12 - 9 = 0$$

디터미넌트가 0으로 계산된 것은 3번째 행에서 2번째 행을 빼어 얻게 되는 $(3,3,3)$이 2번째 행에서 1번째 행을 빼서 얻은 것과 동일하므로 이미 예상되었던 결과이다. 이번에는 두번째 열을 따른 계산을 보였다.

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = -2 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 5 \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} - 8 \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}$$

$$= (-2) \cdot (-6) + 5 \cdot (-12) - 8 \cdot (-6) = 12 - 60 + 48 = 0$$

남은 4개의 계산도 수행하여 같은 결과가 나옴을 확인해 보기 바란다.

크레이머의 규칙(*Cramer's Rule*)이란 $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$일 때 방정식 $AX = B$의 해를

$$x_i = \frac{\det A_{(i)}}{\det A}$$

로 구하는 것을 뜻한다. 단, 여기서 x_i는 X의 i-항을 뜻하며 $A_{(i)}$는 A의 i째 열을 B로 대체하여 얻은 정방행렬을 뜻한다. 이 규칙을 사용하려면 $\det A \neq 0$가 만족되어야 함은 물론이다. 증명은 아래의 연습문제 이후에 하겠다.

<u>연습문제</u> 3.76 [연습문제 3.30]에서 풀었던 4개의 선형연립방정식들을 크레이머 규칙을 사용하여 풀어 보라. ⊣

(증명). of Cramer's Rule: 당분간 i를 고정한다. A의 i-열을 $\vec{x} \stackrel{\text{def}}{=} (x_1, \ldots, x_n)^T$로 대체하여 얻은 행렬의 디터미넌트는 i-열을 따른 여인수 전개에 의하여 $x_{1i}c_{1i} + \cdots + x_{ni}c_{ni}$로 계산된다. $\vec{x}$가 A의 $i' \neq i$째 열일 때는 이 값이 0이 된다.

연립방정식 $AX = B$를 다음과 같이 써 보자.

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m
\end{aligned}$$

각 $j = 1, \ldots, n$에 대하여 이 연립방정식의 j번째 등식의 양변에 c_{ji}를 곱하여 모두 더하여 얻은 등식을 (*EQN*)이라 부르기로 하자. (*EQN*)의 우변은 $\det A_{(i)}$가 되고, 좌변은 적당한 계수 $d_1, \ldots, d_n$에 대하여 $d_1x_1 + \cdots d_nx_n$이 될 것이다.

$d_1 = a_{1i}c_{1i} + \cdots + a_{ni}c_{ni}$의 값은 A의 i-열을 A의 첫째 열로 대체하여 얻은 행렬의 디터미넌트이므로 그 값은 0이다. 물론 $i \neq 1$이라는 전제 하에 이러하다. $d_{i'}$의 값은 같은 이유로 $i' \neq i$일 때는 전부 0이고, $i' = i$일 때는 A의 i-열을 A의 i-열로 대체하여 얻은 행렬, 즉 A의 디터미넌트 $\det A$이다. 따라서 (*EQN*)은 $\det A \cdot x_i = \det A_{(i)}$가 되며, 이것으로부터 우리가 원하던 $x_i = \frac{\det A_{(i)}}{\det A}$를 얻는다. (*Q.E.D.*)

<u>연습문제</u> 3.77 $AA_{\text{cof}} = (\det A)I$를 증명함으로써 $A^{-1} = A_{\text{cof}}/\det A$를 유도하여라. ⊣

4

토픽들

4.1 사영과 근사

4.1.1 선형방정식의 근사해

내적공간 V의 부분공간 U와 $\vec{v} \in V$가 주어졌을 때 U의 원소 중에서 $\vec{v}$에 가장 근사한 것은 $\vec{v}$를 U로 사영하여 얻을 수 있다는 사실을 보조정리 2.40에서 보였다. 그리고 이러한 사영 $\mathrm{Proj}_U\vec{p}$를 얻는 방법은 pp32–34에 설명하였다. 이 내용은 $\dim(U) = 1$인 경우에 대해서는 이에 앞서 [정리 1.13]에 설명되어 있다.

이 방법으로 원하는 근사값을 얻기 위하여는 먼저 직교기저를 찾아야 하며, 이것의 계산에는 적지 않은 시간이 필요하다는 약점이 있을 수 있다. 다음과 같은 문제를 생각해 보자.

> $A \in \mathbb{R}^{m\times n}$, $\vec{y} \in \mathbb{R}^{m\times 1}$에 대해서 $\vec{b}$에 대한 방정식 $A\vec{b} = \vec{y}$의 해가 존재하지 않을 때, $\|A\vec{b} - \vec{y}\|$가 최소가 되도록 하는 $\vec{b} \in \mathbb{R}^{n\times 1}$을 구하여라.

얼핏 보기에는 이 문제가 "사영"과는 별 관련이 없어 보일 수 있지만 이것을 "근사해"를 구하는 문제로 파악하는 순간 사영의 문제로 환원됨을 알 수 있을 것이다.

이 문제는 $V \stackrel{\text{def}}{=} \mathbb{R}^{m\times 1}$, $U \stackrel{\text{def}}{=} \mathrm{ColSpace}(A)$로 두었을 때 주어진 $\vec{y} \in V$의 U에 대한 사영 $\mathrm{Proj}_U\vec{y}$을 구하는 문제로 볼 수 있다.

pp32–34에 기술된 방법을 따르자면 먼저 $\mathrm{ColSpace}(A) = U$의 직교기저를 구해야 한다. 그램-슈미트 과정은 당장 사용할 수 없다. 왜냐하면 이것은 주어진 기저로부터 직교기저를 구하는 방법인데 현재 주어진 것은 U의 기저가 아니라 생성집합(spanning set), 즉 A의 열벡터들 뿐이기 때문이다. 따라서 일단 가우스-조던 소거를 통하여 기저를 구하고 그 다음에 그램-슈미트 과정을 사용해서 직교기저를 구해야 한다. 이런 계산 방법은 시간이 많이 걸릴 뿐만 아니라, 만일 A가 실측 데이터로부터 주어진 행렬이라면 가수스-조던 소거 계산을 하는 중에 0에 아주 가까운 피봇(pivot) 원소(즉, 대각항)이 얻어지게 되면 큰 오차가 발생하는 문제도 있다. 또한 결과로 얻어진 근사값을 닫힌 형식의 수식(closed form formula)으로 표현하기가 무척 어려워 보이는 문제도 있다.

깔끔한 해법을 아래에 보였다. 사영의 정의에 의하여

$$\left(\vec{y} - \mathrm{Proj}_U\vec{y}\right) \perp U \tag{4.1}$$

이어야 한다. (이를 만족하는 $\mathrm{Proj}_U\vec{y}$가 유일하게 존재함은 증명된 바 있다.) $\mathrm{Proj}_U\vec{y}$는 $U = \mathrm{ColSpace}(A)$의 원소이므로 $A\vec{b}$로 쓸 수 있다. 그리고 $\vec{y} - \mathrm{Proj}_U\vec{y}$가 U와 수직이라는 것은 U의 생성집합인 A의 열벡터들과 수직이라는 것과 동등하므로 이제 (4.1)은 다음과 같이 바꿔 쓸 수 있다.

$$A^T\left(\vec{y} - A\vec{b}\right) = \vec{0} \tag{4.2}$$

$$A^T\vec{y} = A^TA\vec{b} \tag{4.3}$$

여기서 만일 A^TA가 가역행렬이라면 다음과 같은 (근사)해를 얻게 된다.

$$\vec{b} = \left(A^TA\right)^{-1}A^T\vec{y} \tag{4.4}$$

4.1.2 회귀(Regression)

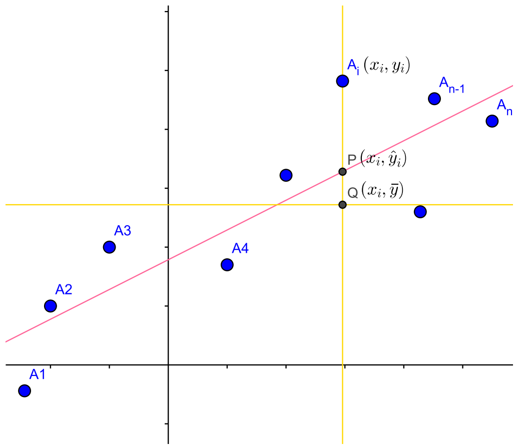
선형방정식의 근사해는 회귀(*regression*) 문제를 푸는 멋진 도구이다. 회귀 중에서 가장 간단한 단순선형회귀(*simple linear regression*) 문제는 다음과 같다.

평면 상에 n개의 점

$$A_1 \stackrel{\mathrm{def}}{=} (x_1, y_1), \ldots A_n \stackrel{\mathrm{def}}{=} (x_n, y_n)$$

가 주어졌다. 이 점들과의 거리의 제곱의 합이 최소가 되는 직선, 즉 적합선의 기울기 b_1과 y-절편 b_2를 찾고자 한다. 오른편 그림에서 적합선을 붉은 직선으로 나타내었다. 적합선의 방정식

$$y = b_1x + b_2$$



에 의하여 계산되는 $x = x_i$에서의 y-좌표를 $\hat{y}_i = b_1x_i + b_2$로 놓으면, 거리의 제곱의 합이 최소가 된다는 것은 $SSE \stackrel{\mathrm{def}}{=} \sum_{i=1}^{n}(\hat{y}_i - y_i)^2$이 최소가 된다는 것과 동등하다.

$SSE = \sum_{i=1}^{n}\left((b_1x_i + b_2) - y_i\right)^2$을 b_1와 b_2의 2변수 함수로 보고 미분에 의하여 적합선의 방정식을 구하는 것이 흔히 쓰이는 방법이지만, 여기서는 사영에 의한 해법

(4.4)를 이용하여 $\vec{b} \stackrel{\text{def}}{=} (b_1, b_2)^T$ 를 구하고자 한다.

$$SSE = \left\| \begin{bmatrix} b_1x_1 + b_2 \\ \vdots \\ b_1x_n + b_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\|^2$$

이므로

$$A \stackrel{\text{def}}{=} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \qquad A^T = \begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \qquad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

로 두면

$$SSE = \left\| A\vec{b} - \vec{y} \right\|^2$$

이 되어

$$\begin{aligned} \vec{b} &= \left(\begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\ &= \begin{bmatrix} \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i & n \end{bmatrix}^{-1} \begin{bmatrix} \sum_i x_i y_i \\ \sum_i y_i \end{bmatrix} \end{aligned} \tag{4.5}$$

를 얻게 된다. 여기서 역행렬의 존재는 코시-슈바르츠 부등식 $\|\vec{u}\|^2 \|\vec{v}\|^2 \geq |\vec{u} \cdot \vec{v}|^2$ 에 의해서 보장 받는다. 즉, $\vec{u} = (x_1, \ldots, x_n)$, $\vec{v} = (1, \ldots, 1)$ 로 두면 $n \sum_i x_i^2 \geq (\sum_i x_i)^2$ 가 성립하며 이때 x_i 들이 모두 같은 값을 갖지 않는 한 $\geq$는 $>$로 바꿔쓸 수 있다.

(4.5)에서 멈추지 않고 더 계산하면 잘 알려진 공식

$$b_1 = \frac{S_{XY}}{S_{XX}}, \tag{4.6}$$

$$b_2 = \bar{y} - b_1\bar{x}, \text{ where} \tag{4.7}$$

$$S_{XY} \stackrel{\text{def}}{=} \textstyle\sum_i (x_i - \bar{x})(y_i - \bar{y}) = \sum_i x_i y_i - n\bar{x}\bar{y},$$

$$S_{XX} = \stackrel{\text{def}}{=} \sum_i (x_i - \bar{x})^2 = \sum_i x_i^2 - n\bar{x}^2,$$

$$\bar{y} \stackrel{\text{def}}{=} \textstyle\sum_i y_i / n, \quad \bar{x} \stackrel{\text{def}}{=} \sum_i x_i / n$$

을 얻을 수 있다. 계산과정을 아래 보였다.

$$\begin{aligned}
\vec{b} &= \frac{1}{n\sum_i x_i^2 - (\sum_i x_i)^2} \begin{bmatrix} n & -\sum_i x_i \\ -\sum_i x_i & \sum_i x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i x_i y_i \\ \sum_i y_i \end{bmatrix} \\
&= \frac{1}{nS_{XX}} \begin{bmatrix} n\sum_i x_i y_i - \sum_i x_i \sum_i y_i \\ -\sum_i x_i \sum_i x_i y_i + \sum_i x_i^2 \sum_i y_i \end{bmatrix} \\
&= \frac{1}{nS_{XX}} \begin{bmatrix} n\sum_i x_i y_i - n^2 \bar{x}\bar{y} \\ -n\bar{x}\sum_i x_i y_i + \sum_i x_i^2 (n\bar{y}) \end{bmatrix} = \frac{1}{S_{XX}} \begin{bmatrix} S_{XY} \\ -\bar{x}\sum_i x_i y_i + \bar{y}\sum_i x_i^2 \end{bmatrix}
\end{aligned}$$

마지막 행렬의 둘째 항을 다음과 같이 고쳐 쓴다.

$$\begin{aligned}
&\left(-\bar{x}(\textstyle\sum_i x_i y_i - n\bar{x}\bar{y}) - n\bar{x}^2\bar{y}\right) + \bar{y}(S_{XX} + n\bar{x}^2) \\
=& -\bar{x}S_{XY} + \bar{y}S_{XX}
\end{aligned}$$

그러면 마침내

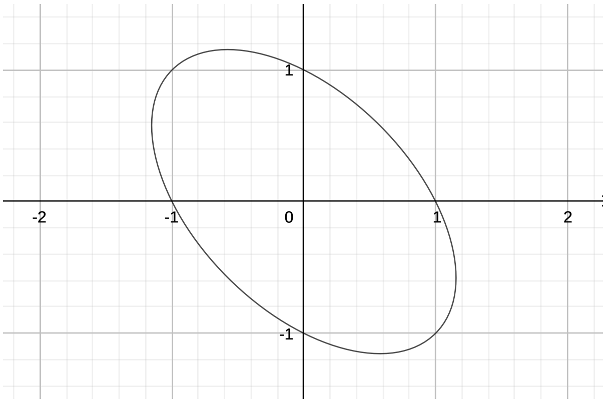
$$\vec{b} = \frac{1}{S_{XX}} \begin{bmatrix} S_{XY} \\ \bar{y}S_{XX} - \bar{x}S_{XY} \end{bmatrix} = \begin{bmatrix} S_{XY}/S_{XX} \\ \bar{y} - (S_{XY}/S_{XX})\bar{x} \end{bmatrix} = \begin{bmatrix} S_{XY}/S_{XX} \\ \bar{y} - b_1\bar{x} \end{bmatrix}$$

의 계산 결과를 얻고 이로부터 우리가 보이려던 (4.6)과 (4.7)을 얻게 된다. □

4.2 좌표축의 회전

평면상에서 $x^2 + xy + y^2 = 1$을 만족하는 점 (x, y)의 집합은 어떤 도형을 이루는가? 실수 a, b, c, d, e에 대해서 $ax^2 + by^2 + cxy + dx + ey = 1$을 만족하는 점 (x, y)의 집합은 어떤 도형을 이루는가? 실수 $a, \ldots, i$에 대하여 $\mathbb{R}^3$에서 $ax^2 + by^2 + cz^2 + dxy + eyz + fyz + gx + hy + iz = 1$을 만족하는 점 (x, y, z)의 집합은 어떤 도형을 이루는가?

위의 질문들에 대한 답은 좌표축의 회전과 밀접한 관계가 있다. 첫번째 질문에 대한 답을 아래에 보였다.



x, y를 아래와 같이 x', y'으로 치환하면

$$x = \frac{1}{\sqrt{2}}(x' - y')$$
$$y = \frac{1}{\sqrt{2}}(x' + y')$$

원래의 방정식 $x^2 + xy + y^2 = 1$은 다음과 같은 타원의 방정식으로 바뀐다.

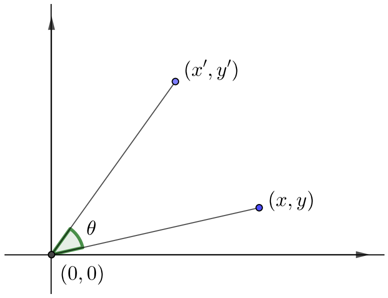
$$\frac{x'^2}{\left(\sqrt{2/3}\right)^2} + \frac{y'^2}{(\sqrt{2})^2} = 1 \tag{4.8}$$

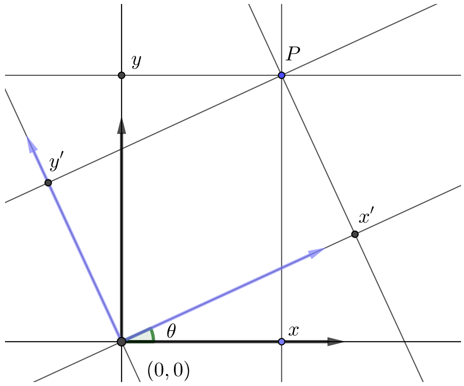
(4.8)은 그 위에 보인 타원의 방정식으로 해석할 수 있음을 설명하고자 한다. 다음의 물음에 답해야 할 것이다: x', y'의 의미는 무엇인가? 그리고 (x, y)와 (x', y')의 관계는 무엇인가?

$\mathbb{R}^2$에서 원점을 중심으로 일정한 각도 θ만큼 회전시키는 것은 선형사상임을 직관적으로 쉽게 알 수 있을 것이다.

아래의 왼쪽 그림은 평면 상에서 좌표가 (x, y)인 점을 θ만큼 회전시켜 얻은 점의 좌표를 (x', y')으로 둔 것이다. $(x', y')^T = M_1(x, y)^T$가 되도록 행렬 $M_1 \in \mathbb{R}^2$을 구해 보자. 참고로 수학에서는 회전의 방향을 $\theta > 0$이면 반시계방향, $\theta < 0$이면 시계방향인 것으로 정해 놓았음을 말해 둔다.

아래의 오른쪽 그림은 점 P의 좌표를 (x, y)라 하였을 때, 좌표계를 θ만큼 회전시켜 얻은 새로운 좌표계에서의 좌표를 (x', y')으로 나타낸 것이다. $(x', y')^T = M_2(x, y)^T$가 되도록 행렬 $M_2 \in \mathbb{R}^2$를 구해 보자.





4.3 직교행렬

Lemma *4.1* 정방행렬 $A \in \mathbb{R}^{n\times n}$ 의 열벡터들이 $\mathbb{R}^n$ 의 단위직교기저를 이루는 것은 A 의 행벡터들이 $\mathbb{R}^n$ 의 단위직교기저를 이루는 것의 필요충분조건이다.

Proof. 간단한 연습문제임. □

Definition 4.2 $A \in \mathbb{R}^{n\times n}$ 은 $A^{-1} = A^T$ 일 때 직교행렬(*orhthogonal matrix, unitary matrix*)이라고 한다. (Unitary matrix라는 용어는 $\mathbb{C}^{n\times n}$ 에 대해서 주로 쓰인다.) ⊣

$A^T = A^{-1}$, 대칭과 회전의 조합으로 이루어진다.

4.4 행렬의 고윳값, 고유벡터, 대각화

$\det(\lambda I - A) = 0$

4.5 행렬의 분해

SVD, LU, Householder transformation and QR-decomposition, Jordan-Normal Form

찾아보기

$\vec{0}$, 2
$\vec{u} \cdot \vec{v}$, 4
$\|\vec{u}\|$, 4
$\mathrm{Proj}_{\vec{q}}\,\vec{p}$, 7
$\det(A)$, 10
$\vec{p} \times \vec{q}$, 14
$\mathrm{Span}(U)$, $\mathrm{Span}(\vec{p}_1, \ldots, \vec{p}_m)$, 24
$\vec{e}_i$, 26
$\dim(V)$, 27
$\mathrm{Proj}_U \vec{p}$, $\mathrm{Proj}_U W$, 32
$\perp$, $\vec{u} \perp \vec{v}$, $\vec{u} \perp U$, $U \perp W$, 37
$U^{\perp}$, 37
$X^{m \times n}$, 39
$(A)_{ij}$, a_{ij}, 39
I, I_n, 41
A^{-1}, 41
O, 42
A^T, A^H, 43
$\vec{x}^T \vec{y}$, $\vec{x}^H \vec{y}$, 43
P_{ij}, 48
$I_{m,i}(\alpha)$, 48
$E_{ij}(\alpha)$, 48
$\mathrm{Ker}(L), \mathrm{Ran}(L)$, 54
$\mathrm{ColSpace}(A)$, $\mathrm{RowSpace}(A)$, 56
$\mathrm{rank}(A)$, 57
$M_B(v)$, 62
$M_{B_2}^{B_1}(L)$, 63
id_V, 63
$\det(A)$, $|A|$, 70
$A_{(ij)}$, 72
A_{cof}, 72

back substitution, 52
basis, 26
Bessel inquality, 36

Cauchy-Schwarz inequality, 5
closed
    under addition, scalar multi., 23
codomain
    of matrix, 39
coefficient, 8, 24
cofactor expansion, 72
cofactor matrix, 72
column, 39
column space
    of a matrix, 57
column vector, 39
complement
    in a vector space, 37
component, 2
    of $\vec{p}$, parallel and normal, 7
coordinate vector, 62
Cramer's Rule, 73
cross product, 14

decomposition of matrices
    LDU, 53
determinant, 14, 44, 70
    of a matrix, 10
diagonal element, 52
diagonal matrix, 43
dimension, 27
dot product, 4

echelon form matrix, 61
elementary matrix, 48
elementary row operation, 47
entry
    of matrix, 39

forward elimination, 52

fundamental vector, 26

Gauss-Jordan elimination, 47, 52
Gaussian elimination, 47
generating set, 24
Gram-Schmidt process, 34

identity matrix, 41
(i, j)-minor
    정방행렬의, 72
inner product, 4
inverse matrix, 41
invertible matrix, 42

kernel, 54

LDU decomposition, 53
LDU 분해, 53
linear combination, 8, 24
linear independence, 12
linear mapping/transformation, 54
lower triangular matrix, 43

matrix, 9, 39
matrix multiplication, 39
multiplication
    matrix, 39

norm, 4
    of a vector, 4
normal, 6
normal vector, 17

ordered basis, 62
orthogonal, 6, 21
orthogonal basis, 34
orthogonal complement, 37
orthonormal basis, 34

parameter, 16
perpendicular, 6, 21
pivot element
    of row vectors, 61
projection
    of $\vec{p}$ on $\vec{q}$, 7
    of $\vec{p}$ on U, 32

range, 54
rank
    of matrix, 57
reduced echelon form matrix, 61
regression
    simple linear, 76
representation matrix
    of a linear mapping, 63
row, 39
row space
    of a matrix, 57
row vector, 39

scalar, 1
scalar multiplication, 1
scalar product, 4
Schwarz inequality, 5
singular matrix, 53
span, 24
square matrix, 39
standard basis, 26
subspace, 23
symmetric matrix, 43

transformation of coordinate system, 63
transpose
    of a matrix, 43
triangle inequality, 5

unit vector, 4
upper triangular matrix, 43

vector
    in $\mathbb{R}^n$, 1
vector space, 1, 21

zero matrix, 42
zero vector, 2

가역행렬, 42
가우스 소거, 47
가우스-조던 소거, 47, 52
계수, 8, 24
공역
　행렬의, 39
그램-슈미트 과정, 34
기본 행작업, 47
기본 행작업 행렬, 48
기본벡터, 14, 26
기약 사다리꼴 행렬, 61
기저, 26
길이
　벡터의, 4

내분점, 9
내적, 4
단위벡터, 4
단위행렬, 41
닫혀있다
　덧셈/스칼라곱에 대하여, 23
대각항, 52
대각행렬, 43
대칭행렬, 43
디터미넌트, 14, 70
　행렬의, 10
랭크
　행렬의, 57
매개변수, 16
법선벡터, 17
벡터
　$\mathbb{R}^n$의 원소, 1
　연산, 1
벡터공간, 1, 21
벡터합
　행렬의, 39
변역, 54
부분공간, 23

사다리꼴 행렬, 61
사영
　벡터에 대한, 7
　부분공간에 대한, 32
상삼각행렬, 43
생성집합, 24
선형독립, 12
선형변환, 선형사상, 54
선형조합, 8, 24
　0 아닌, 24
성분, 2
　벡터의 평행, 수직방향, 7
수직, 21
수직여공간, 37
순서기저, 62
순열행렬, 53
스칼라, 1
스칼라곱, 1, 39
　행렬의, 39
스팬, 24
스팬되다, 스팬하다, 24
여공간, 37
여인수 전개, 72
여인수 행렬, 72
역행렬, 41
열, 39
열공간
　행렬의, 57
열벡터, 39
영벡터, 2
0-선형조합, 24
영행렬, 42
외적, 14
적합선, 76
전치행렬, 43
정규직교기저, 34
정방행렬, 39
좌표계 변환, 63
좌표벡터, 62
직교기저, 34
직교여공간, 37
차원, 27

커늘, 54
코시-슈바르츠 부등식, 5

크기
벡터의, 4
크리이머 규칙, 73
특이행렬, 53
표상행렬
선형사상의, 63
표준기저, 26
피봇 원소
행벡터의, 61
하삼각행렬, 43
항
행렬의, 39
행, 39
행공간
행렬의, 57
행렬, 9, 39
행렬식, 44
행벡터, 39
회귀
단순선형, 76