기초 선형대수학 강의록

김상집

안암 수학강의록 시리즈
제 3권

고려대학교 수학과

기초 선형대수학 강의록

김상집

김상집
고려대학교 수학과
sk23@korea.ac.kr

Mathematics Subject Classification (2010): 15-01
편집 및 발행: 고려대학교 수학과
발행일: 2022년 1월 (v 0.1)
이 책은 TEX과 Oblivoir 클래스를 사용해 조판하였다. 글꼴은 TeX Gyre Pagella, TeX Gyre Heros, Source Han Serif 및 Noto Sans를 사용하였다.

기초 선형대수학 강의록

김상집

차 례

서문

기초 선형대수학 강의록입니다. 두학기 분량으로 구성된 선형대수학 강의 첫번째 학기 내용에 해당합니다. 내용과 구성에 대한 의견 제안 및 오류 제보 등은 이메일 (sk23@korea.ac.kr)로 보내주기 바랍니다.

2022년 1월
김상집

Chapter **1**

Vector space

벡터와 벡터공간을 정의합니다. 벡터공간의 특별한 부분집합으로 부분공간을 정의하고, 부분공간의 원소들을 표현하는 방식에 대해 알아봅시다.

1.1 Definition and examples

Once and for all, we let F denote the set of real numbers or the set of complex numbers.

Definition 1.1.1. *A vector space V over F is a set with addition and scalar multiplication:*

$$V \times V \to V, \quad (v, w) \mapsto v + w$$
$$F \times V \to V, \quad (k, v) \mapsto kv$$

such that

1. *for all $v, w \in V$, $v + w = w + v$*
2. *for all $u, v, w \in V$, $(u + v) + w = u + (v + w)$*
3. *there is a zero vector, denoted by 0_V, in V such that*
$$v + 0_V = 0_V + v = v \quad \text{for all } v \in V.$$
4. *for each $v \in V$, there is an additive inverse w of v in V such that*
$$v + w = w + v = 0_V.$$
5. *for all $a, b \in F$ and $v \in V$, $(a + b)v = av + bv$.*
6. *for all $a \in F$ and $v, w \in V$, $a(v + w) = av + aw$.*
7. *for all $a, b \in F$ and $v \in V$, $(ab)v = a(bv)$.*
8. *for all $v \in V$, $1v = v$.*

Elements in a vector space are called vectors.

Lemma 1.1.2 (Cancellation law). *Let x, y, z be vectors in a vector space V. If*

$$x + z = y + z$$

then $x = y$.

Proposition 1.1.3.

1. *Every vector space V has a unique zero vector 0_V.*
2. *Every vector $v \in V$ has a unique additive inverse.*

Proof. □

We will write $-v$ for the additive inverse of v, and $v - w$ for $v + (-w)$.

Theorem 1.1.4. *Let V be a vector space over F.*

1. $0\,v = v$ *for all* $v \in V$.
2. $a\,0_V = 0_V$ *for all* $a \in F$.
3. $(-a)v = -(av) = a(-v)$ *for all* $a \in F$ *and* $v \in V$.

In particular, for all $v \in V$, the additive inverse $-v$ of v is equal to $(-1)\,v$.

Proof. □

Example 1.1.5 (Vector spaces of matrices)**.** *A $m \times n$ matrix A over F is a rectangular array of numbers in F with m rows and n columns.*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Write $(A)_{ij}$ for the (i, j) entry of A. A $m \times n$ matrix A and a $p \times q$ matrix B are equal, $A = B$, if $m = p$ and $n = q$, and

$$(A)_{ij} = (B)_{ij}$$

for all $1 \le i \le m$ and $1 \le j \le n$.

The set $M_{mn}(F)$ of all $m \times n$ matrices over F with matrix addition $A + B$ and scalar multiplication cA defined by

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij}, \quad (cA)_{ij} = c(A)_{ij}.$$

is a vector space over F. We write $M_n(F)$ for $M_{n,n}(F)$. Its zero vector is the <u>*zero matrix*</u>*, the $m \times n$ matrix all of whose entries are zero, and the additive inverse of A is a $m \times n$ matrix B such that $(B)_{ij} = -(A)_{ij}$ for all i, j.*

Example 1.1.6 (Vector spaces of column vectors). *A $m \times 1$ matrix is called a* <u>*column vector*</u> *of size m and a $1 \times n$ matrix is called a* <u>*row vector*</u> *of size n. The set F^n of all column vectors of size n whose entries are from F with the following operations is a vector space over F.*

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \qquad k \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} kc_1 \\ kc_2 \\ \vdots \\ kc_n \end{bmatrix}.$$

The zero vector is... the additive inverse of ... is ...

Example 1.1.7 (Vector spaces of functions). *The set of functions from a set X to F with function addition $f + g$ and scalar multiplication kf defined by*

$$(f+g)(x) = f(x) + g(x), \quad (kf)(x) = kf(x)$$

is a vector space over F. Then, its zero vector is... the additive inverse of f is ...

Example 1.1.8 (Vector spaces of polynomials). *The set $P^{(m)}$ of polynomials of degree at most m with coefficients from F with polynomial addiion and scalar multiplication*

$$(f+g)(x) = \sum_{i=0}^{m} (a_i + b_i)x^i, \quad (kf)(x) = \sum_{i=0}^{m} (ka_i)x^i$$

is a vector space over F. The zero vector is... the additive inverse of $f \in P^{(m)}$ is ...

1.2 Subspaces

Definition 1.2.1. *A nonempty subset W of a vector space V over F is called a subspace of V, if W is a vector space over F with the same addition and scalar multiplication of V.*

The vector space V itself and $\{0_V\}$ are subspaces of V. The subspace $\{0_V\}$ is called the trivial subspace. We are interested in nontrivial proper subspaces of V

$$\{0_V\} \underset{\neq}{\subsetneq} W \underset{\neq}{\subsetneq} V.$$

Theorem 1.2.2. *Let W be a subset of a vector space V over F. Then, W is a subspace of V if and only if it satisfies the following conditions.*

1. $0_V \in W$.
2. $x + y \in W$ *for all* $x, y \in W$.
3. $kx \in W$ *for all* $k \in F$ *and* $x \in W$.

Proof. □

Example 1.2.3.

1. *For any given $k_1, ..., k_n \in F$, the following set is a subspace of F^n.*

$$W = \{ \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in F^n : \sum_{i=1}^{n} k_i a_i = 0 \}.$$

2. *The subsets of $M_n(F)$ consisting of*
   a) *symmetric matrices: $(A)_{ji} = (A)_{ij}$ for all i and j,*
   b) *skew-symmetric matrices: $(A)_{ji} = -(A)_{ij}$ for all i and j,*
   c) *upper triangular matrices: $(A)_{ij} = 0$ for all $i > j$,*

d) lower triangular matrices: $(A)_{ij} = 0$ *for all* $i < j$,

e) diagonal matrices $(A)_{ij} = 0$ *for all* $i \neq j$

are subspaces of $M_n(F)$.

3. *For any given* $\alpha \in F$, *the following set is a subspace of* $P^{(m)}$.

$$\{ f(x) \in P^{(m)} : f(\alpha) = 0 \}.$$

Proposition 1.2.4. *The intersection of any subspaces of a vector space* V *is a subspace of* V.

Proof. □

1.3 Linear combinations

Definition 1.3.1. *Let V be a vector space over F. A vector $w \in V$ is a linear combination of $v_1, v_2, ..., v_k$, if there are $a_1, ..., a_k \in F$ such that*

$$w = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k.$$

Theorem 1.3.2. *Let S be a nonempty subset of a vector space V over F. The set of all linear combinations of elements in S*

$$\text{Span}(S) = \left\{ \sum_{\textit{finite sum}} a_i v_i : v_i \in S, a_i \in F \right\}$$

is the smallest subspace of V containing S.

The vector space Span(S) is called the subspace of V spanned by S.

Proof. We claim that Span(S) is the intersection of all subspaces of V containing S. Then, by Proposition 1.2.4, it is a subspace of V and any subspace containing S should contain it as a subset. □

Definition 1.3.3. *Let W be a subspace of a vector space V. A spanning set of W is a subset S of W such that*

$$W = \text{Span}(S).$$

Chapter 2

Basis and dimension

벡터공간의 모든 원소를 선형조합 형태로 잘 표현해주는 벡터들의 집합인 기저에 대해 공부합니다. 이를 이용하여 벡터공간의 차원을 정의합니다.

2.1 Linear independence

Definition 2.1.1. *Let V be a vector space over F. Vectors $v_1, v_2, ..., v_k \in V$ are (or the set $\{v_1, ..., v_k\}$ is) linearly dependent, if there are $a_1, a_2, ..., a_k \in F$, not all zero, such that*

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0_V.$$

An infinite subset of V is linearly dependent, if it contains a finite subset that is linearly dependent.

Example 2.1.2. *Let $S = \{v_1, v_2, ..., v_k\} \subset V$.*

1. *If $0_V \in S$ then S is linearly dependent.*
2. *If there is $v_j \in S$ such that*
$$v_j = a_1 v_1 + \cdots + a_{j-1} v_{j-1} + a_{j+1} v_{j+1} + \cdots + a_k v_k$$
for some $a_i \in F$ then S is linearly dependent.

Definition 2.1.3. *Vectors $v_1, v_2, ..., v_k \in V$ are (or the set $\{v_1, ..., v_k\}$ is) linearly independent, if*

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0_V$$

implies that $a_1 = a_2 = \cdots = a_k = 0$. An infinite subset S of V is linearly independent, if every nonempty finite subset of S is linearly independent.

Example 2.1.4.
1. *In the vector space of continuous functions from $[-\pi, \pi]$ to $\mathbb{R}$, the following set is linearly independent.*
$$\{\sin kx, \cos \ell x : 1 \leq k \leq n,\ 0 \leq \ell \leq m\}.$$
2. *In $\mathbb{R}^4$, the following vectors v_1, v_2, v_3, v_4 are linearly dependent*
$$v_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$
while v_1, v_2, v_3 are linearly independent.

Theorem 2.1.5. *Let V be a vector space over F and $S = \{v_1, v_2, ..., v_n\}$ be a subset of V. The set S is linearly independent if and only if every vector in* $\text{Span}(S)$ *can be written as a linear combination of vectors in S in a unique way.*

Proof. □

2.2 Basis

Proposition 2.2.1. *Let V be a vector space over F and $S = \{v_1, ..., v_k\}$ be a subset of V.*

1. *Let S be linearly independent. For any $v \in V \setminus S$, $S \cup \{v\}$ is linearly dependent if and only if $v \in \text{Span}(S)$.*
2. *Let S be a spanning set of V. For any $v \in S$, $S \setminus \{v\}$ is a spanning set of V if and only if $v \in \text{Span}(S \setminus \{v\})$.*

Proof. (1) ($\Rightarrow$) Assume that $S' = S \cup \{v\}$ is linearly dependent. Then, there are coefficients, not all zero, such that

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n + bv = 0_V.$$

We note that b cannot be zero (if $b = 0$ then...). Therefore we have

$$v = (-a_1/b)v_1 + (-a_2/b)v_2 + \cdots + (-a_n/b)v_n.$$

This shows that $v \in \text{Span}(S)$

($\Leftarrow$) Assume that $v \in \text{Span}(S)$. Then,

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

for some coefficients $a_i \in F$. This gives

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n + (-1)v = 0_V,$$

which shows that $S' = \{v_1, ..., v_n, v\}$ is linearly dependent.

(2) Let $S' = S \setminus \{v\}$. □

Definition 2.2.2. *Let V be a vector space over F. A subset $\mathcal{B}$ of V is a* <u>*basis*</u> *of V, if it satisfies the following two conditions:*

1. *every finite subset $\{v_1, ..., v_k\}$ of $\mathcal{B}$ is linearly independent, i.e. if*

$$c_1v_1 + \cdots + c_kv_k = 0_V$$

then $c_1 = \cdots = c_k = 0$.

2. *the set $\mathcal{B}$ spans V, i.e., for every vector $v \in V$ there are $a_1, ..., a_k \in F$ and $v_1, ..., v_k \in \mathcal{B}$ such that*

$$v = a_1 v_1 + \cdots + a_k v_k.$$

Example 2.2.3. *The set $\{E_{ab} : 1 \leq a \leq m,\ 1 \leq b \leq n\}$ is a basis of the vector space $M_{mn}(F)$ where E_{ab} is the $m \times n$ matrix such that*

$$(E_{ab})_{ij} = \begin{cases} 1 & \text{if } i = a \text{ and } j = b, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.2.4. *The elementary basis (or standard basis) for the vector space F^n of column vectors is*

$$\mathcal{E} = \{\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}\}.$$

Example 2.2.5. *The set $B = \{1, x, x^2, ..., x^m\}$ is a basis of the vector space $P^{(m)}$ of polynomials of degree not more than m.*

Let $\mathcal{P}$ be a property we want to investigate. A subset B of a set S is a maximal $\mathcal{P}$ subset of S, if there is no $\mathcal{P}$ subset of S properly containing B, that is, if B' is a $\mathcal{P}$ subset of S and $B \subseteq B'$ then $B' = B$.

Theorem 2.2.6. *Let S be a spanning set of V. Then, a maximal linearly independent subset of S is a basis of V.*

Proof. Let $\mathcal{B}$ be a maximal linearly independent subset of S. To show that it is a basis, it is enough to show that $V = \text{Span}(\mathcal{B})$. Note that $\supseteq$ is trivial. Now we claim

$$S \subseteq \text{Span}(\mathcal{B}).$$

If this is true then, since $\text{Span}(S)$ is the smallest subspace containing S, $\text{Span}(S) \subseteq \text{Span}(\mathcal{B})$. From the hypothesis $V = \text{Span}(S)$,

$$V = \text{Span}(S) \subseteq \text{Span}(\mathcal{B})$$

and we see that $V \subseteq \text{Span}(\mathcal{B})$ and therefore $V = \text{Span}(\mathcal{B})$.

Let us prove the claim. We need to show that for all $v \in S$, $v \in \text{Span}(\mathcal{B})$. (i) if $v \in S$ is an element in $\mathcal{B}$ then it is clear that $v \in \text{Span}(\mathcal{B})$. (ii) if $v \in S \setminus \mathcal{B}$ then $\mathcal{B} \cup \{v\}$ is linearly dependent (because of the maximality condition on $\mathcal{B}$). Thus, we can find some elements $v_i \in \mathcal{B}$ and coefficients, not all zero, such that

$$\sum_{i=1}^{k} a_i v_i + cv = 0_V.$$

In particular, c cannot be zero (why?). Therefore, $v = \sum_{i=1}^{k} (-a_i/c) v_i$ and this shows that $v \in \text{Span}(\mathcal{B})$. □

2.3 Dimension

Theorem 2.3.1 (Replacement theorem). *Let V be a vector space over F and $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis of V. If a subset $S = \{w_1, ..., w_k\}$ of V with k elements is linearly independent then $k \leq n$.*

$$|S| \leq |\mathcal{B}|.$$

Proof. In order to derive a contradiction, suppose that $k > n$. Since $\mathcal{B}$ spans V, $w_1 \in S$ can be written as

$$w_1 = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

with some $a_1, ..., a_n \in F$. Since $w_1 \neq 0_V$, not all a_j are zero. After reindexing them, if necessary, we assume that $a_1 \neq 0$. Then,

$$v_1 = (1/a_1)w_1 - (a_2/a_1)v_2 - \cdots - (a_n/a_1)v_n$$

and we can check that the set obtained from $\mathcal{B}$ by replacing v_1 with w_1

$$\mathcal{B}_1 = \{w_1, v_2, v_3, ..., v_n\}$$

is a basis of V.

In a similar way, we can replace v_2 in $\mathcal{B}_1$ with w_2 to obtain a basis $\mathcal{B}_2$. Continue these procedures until we obtain a basis

$$\mathcal{B}_n = \{w_1, w_2, ..., w_n\}.$$

Then since $\mathcal{B}_n$ spans V, w_{n+1} can be written as a linear combination of $w_1, ..., w_n$, which contradicts to the assumption that S is linearly independent. Therefore, $k \leq n$. □

Theorem 2.3.2. *Let V be a vector space over F and $\mathcal{B}$ be a basis of V having finitely many elements. Then, any other basis $\mathcal{B}'$ of V contains finitely many elements and*

$$|\mathcal{B}'| = |\mathcal{B}|.$$

Proof. Using Theorem 2.3.1, since $\mathcal{B}$ is a basis and $\mathcal{B}'$ is linearly independent, $|\mathcal{B}| \geq |\mathcal{B}'|$. Since $\mathcal{B}'$ is a basis and $\mathcal{B}$ is linearly independent, $|\mathcal{B}| \leq |\mathcal{B}'|$. □

Definition 2.3.3. *The number of vectors in a basis $\mathcal{B}$ of a vector space V is called the dimension of V.*

$$\dim\ V = |\mathcal{B}|$$

Example 2.3.4.

1. *The dimension of $M_{mn}(F)$ is mn.*

2. *The dimension of F^n is n.*

3. *The dimension of $P^{(m)}$ is $m+1$.*

Theorem 2.3.5. *Let W be a subspace of a vector space V.*

1. $\dim W \le \dim V$.

2. *If $\dim W = \dim V$ then $V = W$.*

Proof. (1) Find a basis $\{w_1, ..., w_k\}$ of W. Then, since it is linearly independent in V, $k \le \dim V$ by Theorem 2.3.1.

(2) Let $\dim W = \dim V = n$ and $\mathcal{B}_W$ be a basis of W. Suppose that W is a proper subset of V. Then, we can find $v \in V \setminus W$ so that the following set with $n+1$ elements

$$\mathcal{B}_W \cup \{v\}$$

is linearly independent in V, which contradicts to Theorem 2.3.1. □

Theorem 2.3.6 (Basis extension theorem)**.** *Let V be a vector space with $\dim V = n$ and W be a subspace of V. If $\{v_1, ..., v_k\}$ is a basis of W then there are vectors $v_{k+1}, ..., v_n \in V \setminus W$ such that*

$$\mathcal{B} = \{v_1, v_2, ..., v_k, v_{k+1}, ..., v_n\}$$

is a basis of V.

Proof. From Proposition 2.2.1, if we choose $v_{k+1} \in V \setminus \mathrm{Span}(\{v_1, ..., v_k\})$ then

$$\dim \mathrm{Span}(\{v_1, ..., v_k, v_{k+1}\}) = k + 1.$$

Repeat this procedure until we obtain $W = \mathrm{Span}(\{v_1, ..., v_n\})$. Since W is a subspace of V with dimension equal to $\dim V$, we have $W = V$ and thus $\mathcal{B}$ is a basis of V. □

Chapter 3

Linear transformation

두 벡터공간을 잘 연결해주는 함수 선형변환을 정의하고, 선형변환에 의해 주어지는 정의역과 공역의 부분공간을 공부합니다. 두 벡터공간이 서로 동형이라는 것을 정의하고, 이를 이용하여 유한차원 벡터공간들을 분류합니다.

3.1 Maps between vector spaces

Definition 3.1.1. *Let V and W be vector spaces over the same F. A map T from V to W is a linear transformation, if it satisfies*

$$T(v_1 + v_2) = T(v_1) + T(v_2) \quad \text{and} \quad T(kv) = kT(v)$$

for all $v_1, v_2, v \in V$ and $k \in F$.

Example 3.1.2. 1. $F^2 \to F^2$,

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

2. $P^{(m)} \to P^{(m)}$, $f \mapsto f'$.
3. $P^{(m)}$ over $\mathbb{R} \to \mathbb{R}$, $f \mapsto \int_a^b f(x)\, dx$.
4. The identity map $I_V : V \to V$, $v \mapsto v$.
5. The zero map $T_0 : V \to W$, $v \mapsto 0_W$.
6. $M_{mn}(F) \to M_{nm}(F)$ sending A to its transpose A^T, the $n \times m$ matrix such that
$$(A^T)_{ij} = (A)_{ji} \qquad \text{for } 1 \le i \le m, 1 \le j \le n.$$
7. $M_n(F) \to F$ sending A to its trace $tr(A) = \sum_{i=1}^n (A)_{ii}$.

Proposition 3.1.3. *Let V and W be vector spaces over the same F and $T : V \to W$ be a linear transformation.*

1. $T(0_V) = 0_W$.
2. *For every $v \in V$, the inverse of v in V maps to the inverse of $T(v)$ in W, i.e.,*
$$T(-v) = -T(v) \qquad \text{for all } v \in V.$$
3. $T(\sum_{i=1}^n a_i v_i) = \sum_{i=1}^n a_i T(v_i)$ *for all $a_i \in F$ and $v_i \in V$.*

Proof. □

Proposition 3.1.4. *Let V and W be vector spaces over the same F and $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis of V. For any vectors $w_1, ..., w_n$ in W, there is a unique linear transformation*

$$T : V \to W \text{ such that } T(v_j) = w_j$$

for $j = 1, 2, ..., n$.

Proof. □

Let V and W be vector spaces over the same F. We write $\mathrm{Hom}_F(V, W)$ for the set of all linear transformations from V to W.

Proposition 3.1.5. *The set $\mathrm{Hom}_F(V, W)$ with the following addition and scalar multiplication is a vector space over F. For $T, S \in \mathrm{Hom}_F(V, W)$ and $k \in F$, $T + S$ and kT are the maps from V to W defined by*

$$(T + S)(v) = T(v) + S(v) \text{ and } (kT)(v) = k(T(v)).$$

for all $v \in V$.

Proof. □

Then, the zero vector in $\mathrm{Hom}_F(V, W)$ is the zero map (Example 3.1.2), and the additive inverse of $T \in \mathrm{Hom}_F(V, W)$ is $-T : V \to W$ defined by

$$(-T)(v) = -(T(v)) \text{ for all } v \in V.$$

Proposition 3.1.6. *Let U, V, and W be vector spaces over the same F. If $S : U \to V$ and $T : V \to W$ are linear transformations then their composition*

$$T \circ S : U \to W$$

is also a linear transformation.

Proof. For $u_1, u_2, u \in U$ and $k \in F$,

$$\begin{aligned}
(T \circ S)(u_1 + u_2) &= T(S(u_1 + u_2)) = T(S(u_1) + S(u_2)) \\
&= T(S(u_1)) + T(S(u_2)) = (T \circ S)(u_1) + (T \circ S)(u_2). \\
(T \circ S)(ku) &= T(S(ku)) = T(kS(u)) \\
&= kT(S(u)) = k(T \circ S)(u).
\end{aligned}$$

□

3.2 Subspaces associated with T

Let us recall some terminologies. Let f be a function from X (domain) to Y (codomain). For subsets $A \subseteq X$ and $B \subseteq Y$, the image of A under f and the preimage of B under f are

$$f(A) = \{f(x) \in Y : x \in A\} \text{ and } f^{-1}(B) = \{x \in X : f(x) \in B\}$$

respectively. When $A = X$, we often write $\operatorname{im} f$ for $f(X)$ and call it the range or image of f.

Theorem 3.2.1. *Let $T : V \to W$ be a linear transformation. Then, the kernel and image of T*

$$\ker T = \{v \in V : T(v) = 0_W\} \text{ and } \operatorname{im} T = \{T(v) : v \in V\}$$

are subspaces of V and W respectively.

Proof. (1) We need to show that i) $0_V \in \ker T$, ii) $v_1 + v_2 \in \ker T$ for $v_1, v_2 \in \ker T$, and iii) $kv \in \ker T$ for $v \in \ker T$ and $k \in F$.

(2) We need to show that i) $0_W \in \operatorname{im} T$, ii) $v_1 + v_2 \in \operatorname{im} T$ for $v_1, v_2 \in \operatorname{im} T$, and iii) $kv \in \operatorname{im} T$ for $v \in \operatorname{im} T$ and $k \in F$. □

We remark that the kernel and image of T are also called the null space and range of T, and often denoted by $N(T)$ and $R(T)$ respectively.

Theorem 3.2.2. *Let $T : V \to W$ be a linear transformation. The map T is one-to-one if and only if its kernel is trivial, i.e.,*

$$\ker T = \{0_V\}.$$

Proof. (⇒) Since $T(0_V) = 0_W$ (Proposition 3.1.3), $0_V \in \ker T$. Since T is one-to-one, if $T(v) = 0_W = T(0_V)$ then $v = 0_V$. Thus, $\ker T = \{0_V\}$.

(⇐) Suppose $T(v) = T(v')$. Then, $T(v) - T(v') = 0_W$ and thus $T(v - v') = 0_W$, which implies $v - v' \in \ker T$. Since $\ker T = \{0_V\}$, we have $v - v' = 0_V$ and therefore $v = v'$. □

Proposition 3.2.3. *Let* $T : V \to W$ *be a linear transformation. If* $S = \{v_1, ..., v_n\}$ *spans* V *then*

$$T(S) = \{T(v_1), ..., T(v_n)\}$$

spans $\operatorname{im} T$.

Proof. Need to show that

$$\operatorname{Span}(\{T(v_1), ..., T(v_n)\}) = \operatorname{im} T.$$

($\subseteq$) If w is an element in LHS then it is a linear combination

$$w = a_1 T(v_1) + \cdots + a_n T(v_n)$$

for some $a_1, ..., a_n \in F$. Then, $w = T(\sum a_i v_i)$ and thus it is an element in $\operatorname{im} T$.

($\supseteq$) Let $w \in \operatorname{im} T$. Then, $w = T(v)$ for some $v \in V$ and $v = \sum a_i v_i$. This shows that

$$w = T(v) = T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T(v_i) \in \operatorname{Span}(\{T(v_1), ..., T(v_n)\}).$$

□

3.3 Rank-nullity theorem

The nullity of a linear transformation T is the dimension of the kernel (or null space) of T, and the rank of T is the dimension of the image (or range) of T.

$$\text{Nullity}\, T = \dim \ker T, \qquad \text{Rank}\, T = \dim \text{im}\, T.$$

Theorem 3.3.1 (Rank-Nullity theorem, Dimension theorem). *Let V and W be vector spaces over the same F and $T : V \to W$ be a linear transformation. If the dimension of V is finite then it is the sum of the rank of T and the nullity of T.*

$$\dim V \;=\; \text{Rank}\, T + \text{Nullity}\, T.$$

Proof. We let $\dim V = n$ and $\text{Nullity}\, T = k$. We assume that $0 < k < n$. One can consider the other cases $k = 0$ and $k = n$ separately.

Suppose that $\{v_1, ..., v_k\}$ is a basis of $\ker T$. Then, by Theorem 2.3.6, we can find vectors $v_{k+1}, ..., v_n \in V \setminus \text{Span}(\{v_1, ..., v_k\})$ to form a basis of V

$$\mathcal{B} = \{v_1, v_2, ..., v_k, v_{k+1}, ..., v_n\}.$$

We claim that $S = \{T(v_{k+1}), ..., T(v_n)\}$ is a basis of $\text{im}\, T$.

(i) First, let us show that S spans $\text{im}\, T$. Since $\mathcal{B}$ spans V and $T(v_i) = 0_W$ for $1 \leq i \leq k$, by Proposition 3.2.3 ,

$$\begin{aligned} \text{im}\, T &= \text{Span}(\{T(v_1), ..., T(v_k), T(v_{k+1}), ...T(v_n)\}) \\ &= \text{Span}(\{T(v_{k+1}), ...T(v_n)\}). \end{aligned}$$

(ii) To prove that S is linearly independent, suppose that for some $b_i \in F$,

$$\sum_{i=k+1}^{n} b_i T(v_i) = 0_W.$$

Since T is a linear transformation, we have $T\left(\sum_{i=k+1}^{n} b_i v_i\right) = 0_W$, and thus $\sum_{i=k+1}^{n} b_i v_i \in \ker T$. Since $\{v_1, ..., v_k\}$ is a basis of $\ker T$, there are $c_i \in F$ such that

$$\sum_{i=k+1}^{n} b_i v_i = \sum_{i=1}^{k} c_i v_i$$

and thus

$$(-c_1)v_1 + \cdots + (-c_k)v_k + b_{k+1}v_{k+1} + \cdots + b_n v_n = 0_V.$$

Since $\mathcal{B}$ is a basis, we have $b_i = 0$ for all i. □

Corollary 3.3.2. *Let V and W be finite dimensional vector spaces over the same F and $T : V \to W$ be a linear transformation. If $\dim V = \dim W$ then the following are equivalent.*

1. *T is one-to-one.*
2. *T is onto.*
3. *The rank of T is equal to $\dim V$.*

Proof. □

3.4 Isomorphism

Definition 3.4.1. *Let V and W be vector spaces over the same F.*

1. *A map $T : V \to W$ is an isomorphism, if it is a bijective linear transformation.*

2. *If there is an isomorphism from V to W then we say V is isormorphic to W and write*
$$V \cong W.$$

Example 3.4.2. 1. *For any permutation σ of $\{1, 2, ..., n\}$, the following map is an isomorphism.*

$$T_\sigma : F^n \to F^n, \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} x_{\sigma(1)} \\ x_{\sigma(2)} \\ \vdots \\ x_{\sigma(n)} \end{bmatrix}.$$

2. *The map $M_{mn}(F) \to M_{nm}(F)$ defined by $A \mapsto A^T$ is an isomorphism (Example 3.1.2). In particular, the vector space $F^n = M_{n1}(F)$ of column vectors is isomorphic to the vector space $M_{1n}(F)$ of row vectors.*

Theorem 3.4.3. *Vector space isomorphism is an equivalence relation on any collection of finite dimensional vector spaces over the same F.*

Proof. We need to show that

1. for any V, the identity map $I_V : V \to V$ is an isomorphism.

2. if $T : V \to W$ is an isomorphism then its inverse map $T^{-1} : W \to V$ is also an isomorphism.

3. if $S : U \to V$ and $T : V \to W$ are isomorphisms then their composition $T \circ S$ is an isomorphism from U to W (Proposition 3.1.6).

□

As a consequence of this result, a collection of finite dimensional vector spaces can be partitioned into equivalence classes. We expect that two vector spaces in the same equivalence class share many important properties.

> **Theorem 3.4.4.** *Let V and W are finite dimensional vector spaces over the same F. If $V \cong W$ then* $\dim V = \dim W$.

The converse of Theorem 3.4.4 is also true. See Theorem 4.1.4.

Proof. Let $T : V \to W$ be an isomorphism and $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis of V. We want to show that

$$\{T(v_1), ..., T(v_n)\}$$

is a basis of W.

(1) Since T is surjective, for every $w \in W$ there is $v \in V$ such that $w = T(v)$. Then, $v = \sum_{i=1}^n a_i v_i$ and

$$w = T(v) = T\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i T(v_i).$$

Therefore $\{T(v_1), ..., T(v_n)\}$ spans W (cf. Proposition 3.2.3).

(2) To show that it is linearly independent, suppose $\sum_{i=1}^n a_i T(v_i) = 0_W$. Then, since

$$\sum_{i=1}^n a_i T(v_i) = T(\sum_{i=1}^n a_i v_i),$$

the vector $\sum_{i=1}^n a_i v_i$ is in the kernel of T. Since T is injective, $\ker T = \{0_V\}$ and thus

$$\sum_{i=1}^n a_i v_i = 0_V.$$

This implies that all a_i are zero because $\{v_1, ..., v_n\}$ is linearly independent. □

Chapter 4

Coordinatization

벡터공간의 기저를 이용하여 벡터를 종벡터로 표현할 수 있습니다. 마찬가지로 선형변환은 정의역과 공역의 기저를 이용하여 행렬로 표현할 수 있습니다.

4.1 Coordinate vector of v

Let V be a vector space over F. Once a basis $\mathcal{B} = \{b_1, ..., b_n\}$ of V is given, every $v \in V$ can be expressed as a linear combination of $v_1, ..., v_n$ in a unique way: there are unique $a_1, ..., a_n \in F$ such that

$$v = \sum_{i=1}^{n} a_i v_i.$$

Then the coordinate vector of v with relative to $\mathcal{B}$ is the column vector

$$[v]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}. \tag{4.1.1}$$

Note that the order of elements in $\mathcal{B}$ is important in this context.

Proposition 4.1.1. *Let V be a vector space over F and $\mathcal{B} = \{v_1, v_2, ..., v_n\}$ be a basis of V. The map sending v to its coordinate vector $[v]_{\mathcal{B}}$ relative to a basis $\mathcal{B}$*

$$[\ ]_{\mathcal{B}} : V \longrightarrow F^n, \quad v \mapsto [v]_{\mathcal{B}}$$

is an isomorphism.

Proof. (1) For any element

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in F^n,$$

we have $v = \sum_{i=1}^{n} x_i v_i \in V$ and it satisfies $[v]_{\mathcal{B}} = \mathbf{x}$. This shows that the map $[\ ]_{\mathcal{B}}$ is onto. To show that the map $[\ ]_{\mathcal{B}}$ is one-to-one, for any two vectors $v, v' \in V$ we want to show that $[v]_{\mathcal{B}} = [v']_{\mathcal{B}}$ implies that $v = v'$. Let us write v and v' as linear combinations as

$$v = \sum_{i=1}^{n} a_i v_i \text{ and } v' = \sum_{i=1}^{n} a_i' v_i.$$

If $[v]_{\mathcal{B}} = [v']_{\mathcal{B}}$ then $a_i = a_i'$ for all i and

$$v - v' = v + (-1)v' = \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{n} (-a_i') v_i = \sum_{i=1}^{n} (a_i - a_i') v_i = \sum_{i=1}^{n} 0\, v_i = 0_V.$$

Therefore $v = v'$.

(2) To show that it is a linear transformation, we need to check

$$[v + v']_{\mathcal{B}} = [v]_{\mathcal{B}} + [v']_{\mathcal{B}} \quad \text{and} \quad [kv]_{\mathcal{B}} = k[v]_{\mathcal{B}}$$

for all $v, v' \in V$ and $k \in F$. For $v, v' \in V$, we can find $a_1, ..., a_n, a'_1, ..., a'_n \in F$ such that

$$v = \sum_{i=1}^{n} a_i v_i \text{ and } v' = \sum_{i=1}^{n} a'_i v_i.$$

Then, from $v + v' = (\sum_{i=1}^n a_i v_i) + (\sum_{i=1}^n a'_i v_i) = \sum_{i=1}^n (a_i + a'_i) v_i$,

$$[v + v']_{\mathcal{B}} = \begin{bmatrix} a_1 + a'_1 \\ \vdots \\ a_n + a'_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} a'_1 \\ \vdots \\ a'_n \end{bmatrix} = [v]_{\mathcal{B}} + [v']_{\mathcal{B}}.$$

Also, for every $k \in F$ and $v \in V$, we have $kv = k(\sum_{i=1}^n a_i v_i) = \sum_{i=1}^n (k a_i) v_i$, and thus

$$[kv]_{\mathcal{B}} = \begin{bmatrix} k a_1 \\ \vdots \\ k a_n \end{bmatrix} = k \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = k[v]_{\mathcal{B}}.$$

□

Theorem 4.1.2. *Every n-dimensional vector space V over F is isormorphic to F^n.*

$$V \cong F^n$$

Proof. Choose any basis $\mathcal{B}$ of V. The map $[\ ]_{\mathcal{B}}$ gives an isomorphism from V to F^n. □

Example 4.1.3.

1. $M_{mn}(F) \cong F^{mn}$.

2. $P^{(m)} \cong F^{m+1}$.

3. *The space of $n \times n$ symmetric matrices over F is isomorphic to $F^{n(n+1)/2}$.*

4. *The space of $n \times n$ skew-symmetric matrices over F is isomorphic to $F^{n(n-1)/2}$.*

Theorem 4.1.4. *Let V and W be finite dimensional vector spaces ove the same F. Then,* $\dim V = \dim W$ *if and only if* $V \cong W$.

Proof. ($\Leftarrow$) is given in Theorem 3.4.4. Now we prove ($\Rightarrow$). Let $\dim V = \dim W = n$, and fix bases $\mathcal{B}$ and $\mathcal{C}$ for V and W. We have isomorphisms (Proposition 4.1.1)

$$[\,]_{\mathcal{B}} : V \to F^n, \quad [\,]_{\mathcal{C}} : W \to F^n$$

Then, $[\,]_{\mathcal{C}}^{-1} \circ [\,]_{\mathcal{B}} : V \to W$ is an isomorphism and thus $V \cong W$.

To give a more explicit isomorphism, let $\mathcal{B} = \{v_1, ..., v_n\}$ and $\mathcal{C} = \{w_1, ..., w_n\}$ be bases of V and W respectively. Then, there is a unique linear transformation such that

$$T : V \to W, \quad T(v_j) = w_j \text{ for } 1 \leq j \leq n$$

by Proposition 3.1.4 and it is surjective by Proposition 3.2.3. To show that it is injective, we compute the kernel of T. If $T(v) = 0_W$ for some $v \in V$ and $v = \sum_{j=1}^{n} a_j v_j$ for some $a_j \in F$, then

$$T(v) = T\left(\sum_{j=1}^{n} a_j v_j\right) = \sum_{j=1}^{n} a_j T(v_j) = \sum_{j=1}^{n} a_j w_j = 0_W.$$

Since w_j are linearly independent, $a_j = 0$ for all j and therefore $v = 0_V$. □

4.2 Matrix multiplication

Recall that two matrices are equal, if they are of the same size and their (i, j) entries are equal for all (i, j). We write $(M)_{ij}$ for the (i, j) entry of a matrix M.

Definition 4.2.1 (Matrix multiplication: row by column)**.** *The product AB of a $m \times n$ matrix A and a $n \times \ell$ matrix B is the $m \times \ell$ matrix whose (i, j) entry is*

$$(AB)_{ij} = \sum_{k=1}^{n} (A)_{ik}(B)_{kj}$$

for $1 \leq i \leq m$ and $1 \leq j \leq \ell$.

To compute the (i, j) entry of AB, using the ith row of A and the jth column of B,

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_{ik} b_{kj} \end{bmatrix}.$$

Proposition 4.2.2. *For all $m \times n$ matrices A, $n \times k$ matrices B and B', $k \times \ell$ matrices D, and $c \in F$,*

$$A(B + B') = AB + AB', \quad A(cB) = c(AB), \quad (AB)C = A(BC).$$

Proof. In each case, after verifying that the matrices in the both sides of the equality have the same size, we need to check that for all (s, t),

$$\begin{aligned} (A(B + B'))_{st} &= \sum_{p=1}^{n} (A)_{sp}(B + B')_{pt} \\ &= \sum_{p=1}^{n} (A)_{sp}((B)_{pt} + (B')_{pt}) = \sum_{p=1}^{n} (A)_{sp}(B)_{pt} + \sum_{p=1}^{n} (A)_{sp}(B')_{pt} \\ &= (AB)_{st} + (AB')_{st}. \end{aligned}$$

$$
\begin{aligned}
(A(kB))_{st} &= \sum_{p=1}^{n} (A)_{sp}(kB)_{pt} \\
&= \sum_{p=1}^{n} (A)_{sp}(k(B)_{pt}) = k\sum_{p=1}^{n}(A)_{sp}(B)_{pt} \\
&= k(AB)_{st}
\end{aligned}
$$

$$
\begin{aligned}
((AB)C)_{st} &= \sum_{q=1}^{k}(AB)_{sq}(C)_{qt} = \sum_{q=1}^{k}\left(\sum_{p=1}^{n}(A)_{sp}(B)_{pq}\right)(C)_{qt} \\
&= \sum_{p=1}^{n}(A)_{sp}\left(\sum_{q=1}^{k}(B)_{pq}(C)_{qt}\right) = (A(BC))_{st}.
\end{aligned}
$$

□

A column vector $\mathbf{x} \in F^n$ can be considered a $n \times 1$ matrix. Then, the product of a $m \times n$ matrix $A = (a_{ij})$ and a column vector $\mathbf{x} \in F^n$ is

$$
A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}. \tag{4.2.1}
$$

It is often very useful to notice that the matrix-vector product $A\mathbf{x}$ can be realized as a linear combination of columns of A, and vice versa:

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}
$$

and thus

$$
A\mathbf{x} = x_1A_1 + x_2A_2 + \cdots + x_nA_n \tag{4.2.2}
$$

where $A_k = C_k(A)$ is the kth column of the matrix A. Then, the product of two matrices A and B can be realized as a list of matrix-vector products

$$
AB = \begin{bmatrix} AB_1 & AB_2 & \cdots & AB_\ell \end{bmatrix}
$$

where $B_k = C_k(B)$ is the kth column of B.

Alternatively, we can multiply two matrices column by row.

Proposition 4.2.3 (Matrix multiplication: column by row). *If A is a $m \times n$ matrix and B is a $n \times \ell$ matrix then*

$$\begin{aligned} AB &= \sum_{k=1}^{n} C_k(A)R_k(B) \\ &= \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1\ell} \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \begin{bmatrix} b_{n1} & b_{n2} & \cdots & b_{n\ell} \end{bmatrix} \end{aligned}$$

where $C_k(A)$ is the kth column of A and $R_k(B)$ is the kth row of B.

Proof. Note that for each k, $C_k(A)R_k(B)$ is an $m \times \ell$ matrix and its (i, j) entry is $(A)_{ik}(B)_{kj}$. Then their sum over $k = 1, 2, ..., n$

$$(A)_{i1}(B)_{1j} + (A)_{12}(B)_{2j} + \cdots + (A)_{in}(B)_{nj}$$

is equal to $(AB)_{ij}$ in the definition of AB. □

Example 4.2.4. *The identity matrix I_n is the $n \times n$ matrix such that*

$$(I_n)_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for all $m \times n$ matrices X and $n \times k$ matrices Y,

$$XI_n = X \quad \text{and} \quad I_nY = Y.$$

Definition 4.2.5. *A $n \times n$ matrix A is invertible, if there exists a $n \times n$ matrix B such that*

$$BA = I_n \text{ and } AB = I_n.$$

We write A^{-1} for B and call it the inverse of A.

We will study the properties of invertible matrices later, including the uniqueness of A^{-1} when A is invertible.

4.3 Linear transformations and matrices

4.3.1 Matrix representation of T

Lemma 4.3.1. *Let A be a $m \times n$ matrix over F. Then, the map L_A defined by matrix vector multiplication as in* (4.2.1)

$$L_A : F^n \to F^m, \quad L_A(\mathbf{x}) = A\mathbf{x}$$

is a linear transformation.

Proof. This follows from Proposition 4.2.2. □

Lemma 4.3.2. *Let A and B be $m \times n$ matrices over F. Then, $L_A = L_B$ if and only if $A = B$. That is,*

$$A\mathbf{x} = B\mathbf{x} \quad \text{for all } \mathbf{x} \in F^n$$

if and only if $A = B$ as a matrix.

Proof. (⇒) Note that $L_A(\mathbf{e}_j) = A\mathbf{e}_j$ is the jth column of A. Thus, $C_j(A) = C_j(B)$ for all j. □

Lemma 4.3.3. *For every linear transformation $T : F^n \to F^m$, there is a unique $m \times n$ matrix A such that $T = L_A$, and thus*

$$T(\mathbf{x}) = A\mathbf{x} \quad \text{for all } \mathbf{x} \in F^n.$$

Proof. First we want to find a matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in F^n$. Then the uniqueness of A follows from Lemma 4.3.2.

Using the elementary basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ of F^n and the fact that T is a linear transformation,

$$T(\mathbf{x}) = T\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) = \sum_{i=1}^{n} x_i T(\mathbf{e}_i)$$
$$= \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

For the third equality we used the observation (4.2.2). Therefore, $T = L_A$ where

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}.$$

□

Theorem 4.3.4. *Let V and W be vector spaces over the same F. Let $\mathcal{B} = \{v_1, ..., v_n\}$ and $\mathcal{C} = \{w_1, ..., w_m\}$ be bases of V and W respectively. For every linear transformation $T : V \to W$, there exists a unique $m \times n$ matrix $[T]_{\mathcal{BC}}$ such that*

$$[T(v)]_{\mathcal{C}} \;=\; [T]_{\mathcal{BC}}\, [v]_{\mathcal{B}} \quad \text{for all } v \in V. \tag{4.3.1}$$

The matrix $[T]_{\mathcal{BC}}$ is called the matrix representation of T with relative to bases $\mathcal{B}$ and $\mathcal{C}$,

If $V = W$ and the same basis $\mathcal{B}$ is used for the domain and codomain of T, then we write $[T]_{\mathcal{B}}$ for $[T]_{\mathcal{BB}}$.

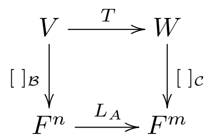
Proof. Note that we are looking for a matrix $A = [T]_{\mathcal{BC}}$ making the diagram in Figure 4.1 commutes. Since $[\,]_{\mathcal{B}}$ is an isomorphism (Proposition 4.1.1), its inverse is an isomorphism from F^n to V and thus

$$[\,]_{\mathcal{C}} \circ T \circ [\,]_{\mathcal{B}}^{-1} : F^n \to F^m$$

is a linear transformation. By Lemma 4.3.3, it should be of the form L_A for some matrix A.

$$[\,]_{\mathcal{C}} \circ T \circ [\,]_{\mathcal{B}}^{-1} \;=\; L_A$$

Therefore, $[T(v)]_{\mathcal{C}} \;=\; L_A[v]_{\mathcal{B}}$ for all $v \in V$. □



⟨**Figure 4.1**⟩ The matrix representation of T with relative to $\mathcal{B}$ and $\mathcal{C}$.

Now we compute the matrix representation $[T]_{\mathcal{BC}}$ explicitly. For every $v \in V$, we can write $v = \sum_{i=1}^{n} a_i v_i$ and

$$[T(v)]_{\mathcal{C}} = \left[T\left(\sum_{i=1}^{n} a_i v_i \right) \right]_{\mathcal{C}} = \sum_{i=1}^{n} a_i [T(v_i)]_{\mathcal{C}}$$

$$= \begin{bmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{C}} & [T(v_2)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & | & & | \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{C}} & [T(v_2)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & | & & | \end{bmatrix} [v]_{\mathcal{B}}.$$

Here, we used the fact that the coordinate map $[\ ]_{\mathcal{C}}$ is a linear transformation (Proposition 4.1.1) for the second equality and the observation (4.2.2) for the third equality.

Proposition 4.3.5. *With the above notation, the matrix representation of $T : V \to W$ with relative to $\mathcal{B}$ and $\mathcal{C}$ is*

$$[T]_{\mathcal{BC}} = \begin{bmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{C}} & [T(v_2)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & | & & | \end{bmatrix}.$$

Example 4.3.6. *Find the matrix representations of*

1. $T : P^{(4)} \to P^{(3)}$, $f \mapsto 3f'$
2. $T : P^{(3)} \to P^{(4)}$, $f \mapsto (2x+1)f$

with relative to bases $\mathcal{B} = \{1, x, x^2, x^3\}$ *for* $P^{(3)}$ *and* $\mathcal{C} = \{1, x, x^2, x^3, x^4\}$ *for* $P^{(4)}$.

Example 4.3.7. *Let* $T : P^{(3)} \to P^{(3)}$ *given by* $f(x) \mapsto xf'(x)$. *Compute* $[T]_{\mathcal{B}}$ *and* $[T]_{\mathcal{B}'}$ *where*

1. $\mathcal{B} = \{4x, 3x^2, 2, x^3\}$.
2. $\mathcal{B}' = \{1 - x, x + x^2, x^2 - x^3, x^3\}$.

Example 4.3.8. *Let* V *and* W *be vector spaces over* F *of dimension* n *and* m *respectively.*

1. *The matrix representation* $[T]_{\mathcal{B}\mathcal{C}}$ *of the zero map* $T_0 : V \to W$ *with relative to any bases* $\mathcal{B}$ *and* $\mathcal{C}$ *is the* $m \times n$ *zero matrix.*
2. *The matrix representation* $[I_V]_{\mathcal{B}}$ *of the identity map* $I_V : V \to V$ *with relative to any basis* $\mathcal{B}$ *for* V *is the identity matrix* I_n.

4.3.2 $\mathrm{Hom}_F(V, W)$ and $M_{mn}(F)$

Recall that $\mathrm{Hom}_F(V, W)$ is the vector space of linear transformations from V to W (Proposition 3.1.5).

Lemma 4.3.9. *Let* V *and* W *be vector spaces over the same* F *with bases* $\mathcal{B}$ *and* $\mathcal{C}$ *respectively. For all linear transformations* $T, S : V \to W$ *and* $k \in F$,

$$[T + S]_{\mathcal{B}\mathcal{C}} = [T]_{\mathcal{B}\mathcal{C}} + [S]_{\mathcal{B}\mathcal{C}} \quad \text{and} \quad [kT]_{\mathcal{B}\mathcal{C}} = k[T]_{\mathcal{B}\mathcal{C}}.$$

Proof. Let $\mathcal{B} = \{v_1, ..., v_n\}$. From Proposition 4.3.5, the jth column of $[T + S]_{\mathcal{B}\mathcal{C}}$ is

$$[(T + S)(v_j)]_{\mathcal{C}} = [T(v_j) + S(v_j)]_{\mathcal{C}} = [T(v_j)]_{\mathcal{C}} + [S(v_j)]_{\mathcal{C}},$$

which is the jth column of $[T]_{\mathcal{B}\mathcal{C}} + [S]_{\mathcal{B}\mathcal{C}}$. The jth column of $[kT]_{\mathcal{B}\mathcal{C}}$ is

$$[(kT)(v_j)]_{\mathcal{C}} = [kT(v_j)]_{\mathcal{C}} = k[T(v_j)]_{\mathcal{C}},$$

which is the jth column of $k[T]_{\mathcal{B}\mathcal{C}}$. □

In Theorem 4.1.2, we saw that every abstract n-dimensional vector over F is isomorphic to F^n. Now we have a similar result for the vector space of linear transformations.

Theorem 4.3.10. *Let V be a vector space of dimension n and W be a vector space of dimension m over the same F. Then,*

$$\mathrm{Hom}_F(V,W) \;\cong\; M_{mn}(F).$$

Proof. Let us fix bases $\mathcal{B}$ and $\mathcal{C}$ for V and W respectively. Then, by Lemma 4.3.9,

$$T \mapsto [T]_{\mathcal{B}\mathcal{C}}$$

is a linear transformation from $\mathrm{Hom}_F(V,W)$ to $M_{mn}(F)$. Now we need to show that this map is one-to-one and onto. □

Example 4.3.11. *In Example 2.2.3, we found the following basis for the vector space $M_{mn}(F)$.*

$$\{E_{a,b} \in M_{mn}(F) : 1 \le a \le m,\ 1 \le b \le n\}.$$

Verify that the elements in $\mathrm{Hom}_F(V,W)$ corresponding to E_{ab} under the above isomorphism are

$$T_{ab} : V \to W, \qquad T_{ab}(v_k) = \begin{cases} w_a & \text{if } b = k, \\ 0 & \text{otherwise.} \end{cases}$$

where $\mathcal{B} = \{v_1, ..., v_n\}$ and $\mathcal{C} = \{w_1, ..., w_m\}$ are bases of V and W respectively, and therefore the maps T_{ab} form a basis of $\mathrm{Hom}_F(V,W)$. See the proof of Theorem 3.4.4.

Next we want to show that matrix multiplication studied in §4.2 can be realized as the composition of linear transformations.

Theorem 4.3.12. *Let U, V, and W be vector spaces over the same F with bases $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ respectively. If $S : U \to V$ and $T : V \to W$ are linear transformations then their composition $T \circ S : U \to W$ is a linear transformation and*

$$[T \circ S]_{\mathcal{A}\mathcal{C}} = [T]_{\mathcal{B}\mathcal{C}}\,[S]_{\mathcal{A}\mathcal{B}}.$$

Thus, the isomorphism given in Theorem 4.3.10 extends to the correspondence between the composition of linear transformations

$$\mathrm{Hom}_F(U,V) \times \mathrm{Hom}_F(V,W) \to \mathrm{Hom}_F(U,W), \quad (S,T) \mapsto T \circ S$$

and matrix multiplication

$$M_{n\ell}(F) \times M_{mn}(F) \to M_{m\ell}(F), \quad (X,Y) \mapsto YX.$$

Proof. The fact that $T \circ S$ is a linear transformation is shown in Proposition 3.1.6. Now, we note that for all $u \in U$,

$$\begin{aligned}([T]_{\mathcal{BC}}\,[S]_{\mathcal{AB}})\,[u]_{\mathcal{A}} &= [T]_{\mathcal{BC}}\,([S]_{\mathcal{AB}}[u]_{\mathcal{A}}) \\ &= [T]_{\mathcal{BC}}[S(u)]_{\mathcal{B}} \\ &= [T(S(u))]_{\mathcal{C}} = [(T \circ S)(u)]_{\mathcal{C}} \\ &= [T \circ S]_{\mathcal{AC}}[u]_{\mathcal{A}}\end{aligned}$$

where we used the associativity of matrix multiplication for the first equality and (4.3.1) for the others. Next, let $\dim U = n$. Since $[\,]_{\mathcal{A}} : U \to F^n$ is surjective (Proposition 4.1.1), the above identities become

$$([T]_{\mathcal{BC}}\,[S]_{\mathcal{AB}})\,\mathbf{x} = [(T \circ S)]_{\mathcal{AC}}\,\mathbf{x} \quad \text{for all } \mathbf{x} \in F^n.$$

Now by applying Lemma 4.3.2 we conclude that $[T]_{\mathcal{BC}}\,[S]_{\mathcal{AB}} = [T \circ S]_{\mathcal{AC}}$ as a matrix. □

Let V be a n-dimensional vector space over F. Linear transformations from V to itself are called endomorphisms of V and we write $\mathrm{End}_F(V)$ for the vector space of endomorphisms of V.

$$\mathrm{End}_F(V) = \mathrm{Hom}_F(V, V) \cong M_n(F).$$

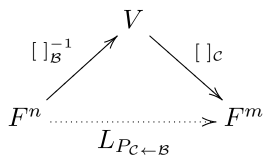
Note that $\mathrm{End}_F(V)$ is closed under composition.

Corollary 4.3.13. *Let V be a finite dimensional vector space over F and $\mathcal{B}$ be a basis of V. For all $T, S \in \mathrm{End}_F(V)$ and $k \in F$,*

$$[T + S]_{\mathcal{B}} = [T]_{\mathcal{B}} + [S]_{\mathcal{B}}, \quad [kT]_{\mathcal{B}} = k[T]_{\mathcal{B}}$$

and

$$[T \circ S]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[S]_{\mathcal{B}}.$$



⟨Figure 4.2⟩ A basis-change matrix.

4.4 Change of basis

4.4.1 Basis change matrix

Let $\mathcal{B}$ and $\mathcal{C}$ be bases of a vector space V over F. Since the coordinate maps $[\,]_{\mathcal{B}}$ and $[\,]_{\mathcal{C}}$ from V to F^n are isomorphisms, so is their composition

$$[\,]_{\mathcal{C}} \circ [\,]_{\mathcal{B}}^{-1} : F^n \to F^n.$$

By Lemma 4.3.3, this map should be of the form L_A for some matrix A, which we will denote by $P_{\mathcal{C}\leftarrow\mathcal{B}}$. See Figure 4.2.

Definition 4.4.1. *With the above notation, the matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$ satisfying*

$$[v]_{\mathcal{C}} \;=\; P_{\mathcal{C}\leftarrow\mathcal{B}}\,[v]_{\mathcal{B}} \quad \text{for all } v \in V$$

is the basis-change matrix (or transition matrix) from $\mathcal{B}$ to $\mathcal{C}$.

Let $\mathcal{B} = \{v_1, \dots, v_n\}$ and compute the matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$. For any $v \in V$, we have $v = \sum_{i=1}^{n} k_i v_i$ and then

$$[v]_{\mathcal{C}} = \left[\sum_{i=1}^{n} k_i v_i\right]_{\mathcal{C}} = \sum_{i=1}^{n} k_i [v_i]_{\mathcal{C}} = \begin{bmatrix} | & & | \\ [v_1]_{\mathcal{C}} & \cdots & [v_n]_{\mathcal{C}} \\ | & & | \end{bmatrix} \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ [v_1]_{\mathcal{C}} & \cdots & [v_n]_{\mathcal{C}} \\ | & & | \end{bmatrix} [v]_{\mathcal{B}}.$$

Proposition 4.4.2. *Let $\mathcal{B} = \{v_1, ..., v_n\}$ and $\mathcal{C}$ be bases of a vector space V. Then the basis-change matrix from $\mathcal{B}$ to $\mathcal{C}$ is*

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} | & & | \\ [v_1]_{\mathcal{C}} & \cdots & [v_n]_{\mathcal{C}} \\ | & & | \end{bmatrix}.$$

Example 4.4.3. *Let $V = P^{(3)}$ and $\mathcal{B} = \{1, x, x^2, x^3\}$. For each of the following cases, find the basis change matrix from $\mathcal{B}$ to $\mathcal{C}$.*

1. $\mathcal{C} = \{1, (x-1), (x-1)^2, (x-1)^3\}$.
2. $\mathcal{C} = \{1, (x+1), (x+1)x, (x+1)x(x-1)\}$.

We note that since $[\ ]_{\mathcal{C}} \circ [\ ]_{\mathcal{B}}^{-1}$ is an isomorphism, the linear transformation $L_{P_{\mathcal{C}\leftarrow\mathcal{B}}}$ is invertible and its inverse map should be

$$L_{P_{\mathcal{C}\leftarrow\mathcal{B}}}^{-1} = L_{P_{\mathcal{B}\leftarrow\mathcal{C}}}.$$

In particular, we have

$$(P_{\mathcal{B}\leftarrow\mathcal{C}} P_{\mathcal{C}\leftarrow\mathcal{B}})\,[v]_{\mathcal{B}} = [v]_{\mathcal{B}} \quad \text{and} \quad (P_{\mathcal{C}\leftarrow\mathcal{B}} P_{\mathcal{B}\leftarrow\mathcal{C}})\,[v]_{\mathcal{C}} = [v]_{\mathcal{C}}$$

for all $v \in V$. See Figure 4.3. Then, by Lemma 4.3.2, we have

$$P_{\mathcal{B}\leftarrow\mathcal{C}} P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{B}} P_{\mathcal{B}\leftarrow\mathcal{C}} = I_n,$$

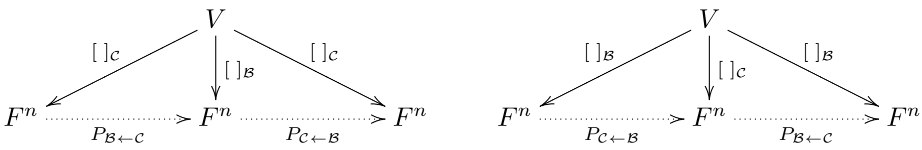
Proposition 4.4.4. *For any two bases $\mathcal{B}$ and $\mathcal{C}$ of a vector space V,*

$$P_{\mathcal{C}\leftarrow\mathcal{B}}^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}} \quad \text{and} \quad P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1} = P_{\mathcal{C}\leftarrow\mathcal{B}}.$$

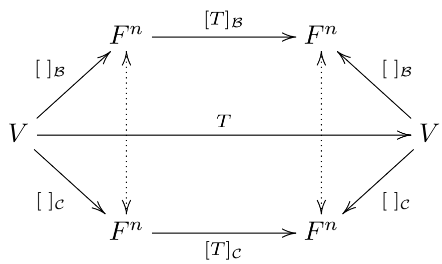
4.4.2 Matrix similarity

Now we focus on a linear transformation $T : V \to V$ with the same basis $\mathcal{B}$ for the domain and codomain. This gives a matrix representation $[T]_{\mathcal{B}}$ of T.

Proposition 4.4.5. *Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Two matrix representations $[T]_{\mathcal{B}}$*



⟨**Figure 4.3**⟩ Basis-change matrices and their inverses.



⟨**Figure 4.4**⟩ Matrix representations $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$.

and $[T]_{\mathcal{C}}$ *of the same linear transformation* $T : V \to V$ *satisfy*

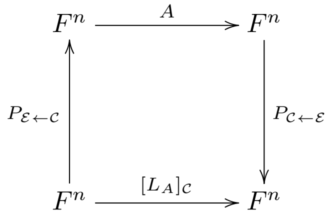
$$[T]_{\mathcal{C}} \;=\; P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,[T]_{\mathcal{B}}\,P_{\mathcal{B}\leftarrow\mathcal{C}}.$$

Proof. See Figure 4.4. For all $v \in V$,

$$\begin{aligned}\left(P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,[T]_{\mathcal{B}}\,P_{\mathcal{B}\leftarrow\mathcal{C}}\right)\,[v]_{\mathcal{C}} &= P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,[T]_{\mathcal{B}}\,\left(P_{\mathcal{B}\leftarrow\mathcal{C}}\,[v]_{\mathcal{C}}\right)\\ &= P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,\left([T]_{\mathcal{B}}\,[v]_{\mathcal{B}}\right) = P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,[T(v)]_{\mathcal{B}}\\ &= P_{\mathcal{C}\leftarrow\mathcal{B}}\,[T(v)]_{\mathcal{B}} = [T(v)]_{\mathcal{C}} = [T]_{\mathcal{C}}\,[v]_{\mathcal{C}}.\end{aligned}$$

Since $[\ ]_{\mathcal{C}} : V \to F^n$ is surjective, we have $\left(P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,[T]_{\mathcal{B}}\,P_{\mathcal{B}\leftarrow\mathcal{C}}\right)\mathbf{x} = [T]_{\mathcal{C}}\,\mathbf{x}$ for all $\mathbf{x} \in F^n$ and therefore, as a matrix, $[T]_{\mathcal{C}} \;=\; P_{\mathcal{B}\leftarrow\mathcal{C}}^{-1}\,[T]_{\mathcal{B}}\,P_{\mathcal{B}\leftarrow\mathcal{C}}$. □

Example 4.4.6. *For* $T : V \to V$ *and bases* $\mathcal{B}$ *and* $\mathcal{B}'$ *of* V *given in Example 4.3.7, find the basis change matrices* $P_{\mathcal{B}\leftarrow\mathcal{B}'}$ *and* $P_{\mathcal{B}'\leftarrow\mathcal{B}}$*, and then verify Proposition 4.4.5.*



⟨**Figure 4.5**⟩ L_A and $[L_A]_{\mathcal{C}}$.

When working with the vector space F^n of column vectors and linear transformations from F^n to itself, we implicitly use the elementary basis $\mathcal{E} = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$ (Example 2.2.4). Note that

1. $\mathbf{v} = [\mathbf{v}]_{\mathcal{E}}$ for all $\mathbf{v} \in F^n$,
2. $A = [L_A]_{\mathcal{E}}$ for any $n \times n$ matrix A.

Now let us compute the matrix representation of $L_A : F^n \to F^n$ with relative to a new basis $\mathcal{C}$ of F^n.

Proposition 4.4.7. *Let $\mathcal{C} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a basis of F^n. The matrix representation of L_A with relative to $\mathcal{C}$ is*

$$[L_A]_{\mathcal{C}} \;=\; P_{\mathcal{E}\leftarrow\mathcal{C}}^{-1}\, A\, P_{\mathcal{E}\leftarrow\mathcal{C}}$$

where

$$P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}.$$

Proof. See Figure 4.5. Note that if $\mathbf{v} \in F^n$ then $[\mathbf{v}]_{\mathcal{E}} = \mathbf{v}$. Then, using Proposition 4.4.2, we have

$$P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} | & | & & | \\ [\mathbf{v}_1]_{\mathcal{E}} & [\mathbf{v}_2]_{\mathcal{E}} & \cdots & [\mathbf{v}_n]_{\mathcal{E}} \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}.$$

□

Definition 4.4.8. *A $n \times n$ matrix A is* similar *to A', if there is an invertible matrix X such that*

$$A' = X^{-1} A X.$$

Proposition 4.4.9. *Matrix similarity is an equivalence relation on $M_n(F)$.*

Proof. (1) For any $A \in M_n(F)$, $A = I_n^{-1} A I_n$ and thus A is similar to itself. (2) Suppose that A is similar to B. Then, there is X such that $B = X^{-1}AX$, and with $Z = X^{-1}$ we have $A = Z^{-1}BZ$, which shows that A is similar to B. (3) Suppose that A is similar to B and B is similar to C. Then, there exisit X and Y such that $B = X^{-1}AX$ and $C = Y^{-1}BY$. With $Z = XY$, we have $C = Z^{-1}AZ$ and therefore, A is similar to C. □

Let $\mathcal{B}$ and $\mathcal{B}'$ be bases of V. Then, the matrix representations $A = [T]_{\mathcal{B}}$ and $A' = [T]_{\mathcal{B}'}$ of the same linear transformation $T : V \to V$ are similar, since $A' = X^{-1}AX$ with

$$X = P_{\mathcal{B} \leftarrow \mathcal{B}'}.$$

On the other hand, the matrix A can be thought of the matrix representation of $L_A : F^n \to F^n$ with relative to the elementary basis $\mathcal{E}$ of F^n. If $\mathcal{C} = \{\mathbf{w}_1, ..., \mathbf{w}_n\}$ is the basis of F^n consisting of the columns of the matrix X (cf. Theorem 6.1.4) then X is equal to the basis change matrix

$$X = P_{\mathcal{E} \leftarrow \mathcal{C}} = \begin{bmatrix} | & & | \\ \mathbf{w}_1 & \cdots & \mathbf{w}_n \\ | & & | \end{bmatrix}.$$

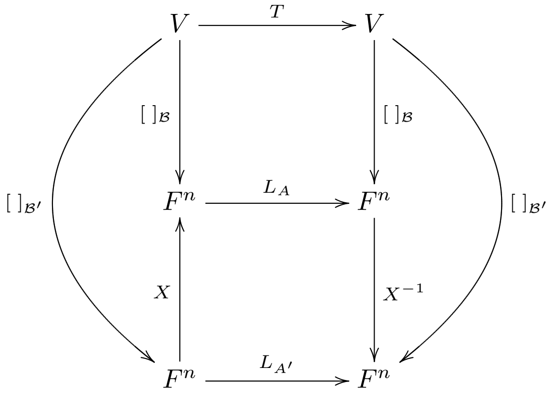
Therefore, $A' = X^{-1}AX$ can be thought of the matrix representation of L_A

$$A' = [L_A]_{\mathcal{C}}$$

with relative to $\mathcal{C}$. See Figure 4.6.

Example 4.4.10. *Consider $T : P^{(3)} \to P^{(3)}$ given by $f(x) \mapsto xf'(x)$. In Example 4.3.7, we computed the matrix representations*

$$A = [T]_{\mathcal{B}} \quad \text{and} \quad A' = [T]_{\mathcal{B}'}$$



〈**Figure 4.6**〉 Similar matrices A and A'.

with relative to the bases $\mathcal{B} = \{4x, 3x^2, 2, x^3\}$ *and* $\mathcal{B}' = \{1 - x, x + x^2, x^2 - x^3, x^3\}$ *of* $P^{(3)}$*. See also Example 4.4.6. Find a basis* $\mathcal{C}$ *of* F^n *such that*

$$A' = [L_A]_{\mathcal{C}}.$$

Chapter 5

Systems of linear equations

미지수 n개와 일차식 m개로 구성된 연립방정식을 풀어봅시다. 방정식의 해를 종벡터의 선형조합으로 표현하고, 이를 선형변환과 연관지어 공부해봅시다.

5.1 Gaussian elimination

A system of m linear equations with n unknowns is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

It is consistent, if it has at least one solution; it is inconsistent, if it does not have any solution.

Using matrix multiplication, we can write the system as $A\mathbf{x} = \mathbf{b}$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Then the augmented matrix for the system is

$$[A|\mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Definition 5.1.1. *In a matrix, if there is a row containing a nonzero entry then the leftmost nonzero entry is called the leading coefficient (or pivot) of that row. A matrix is said to be in reduced row echelon form, if*

1. *all of the leading coefficients are equal to* 1.
2. *the leading coefficient in each row is to the right of the leading coefficient of the row above.*
3. *in every column containing a leading coefficient, all of the other entries in that column are zero.*

The first two conditions imply that the lower left part of the matrix consists of zeros, and the zero rows are located at the bottom of the matrix.

There are three types of elementary row operations we can apply to a matrix:

1. Exchange R_i and R_j:

$$\begin{bmatrix} - & R_i & - \\ - & R_j & - \end{bmatrix} \rightsquigarrow \begin{bmatrix} - & R_j & - \\ - & R_i & - \end{bmatrix}$$

2. Replace R_j with kR_j:

$$\begin{bmatrix} - & R_i & - \end{bmatrix} \rightsquigarrow \begin{bmatrix} - & kR_i & - \end{bmatrix}$$

3. Replace R_j with $kR_i + R_j$:

$$\begin{bmatrix} - & R_i & - \\ - & R_j & - \end{bmatrix} \rightsquigarrow \begin{bmatrix} - & R_i & - \\ - & kR_i + R_j & - \end{bmatrix}$$

To solve a system $A\mathbf{x} = \mathbf{b}$,

1. apply elementary row operations to the augmented matrix for the system

$$[A|\mathbf{b}] = \left[\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & b_3 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & b_4 \end{array}\right]$$

2. to obtain $[\text{rref}(A)|\mathbf{b}']$
$$[\text{rref}(A)|\mathbf{b}'] = \left[\begin{array}{cccccc|c} 0 & 1 & * & 0 & 0 & * & b_1' \\ 0 & 0 & 0 & 1 & 0 & * & b_2' \\ 0 & 0 & 0 & 0 & 1 & * & b_3' \\ 0 & 0 & 0 & 0 & 0 & 0 & b_4' \end{array}\right].$$
3. Note that elementary row operations do not change the solution set of the associated system, i.e., two systems $A\mathbf{x} = \mathbf{b}$ and $\text{rref}(A)\mathbf{x} = \mathbf{b}'$ have the same set of solutions.
4. Solve the simplified system $\text{rref}(A)\,\mathbf{x} = \mathbf{b}'$. Write solutions using free variables for x_j corresponding to columns without leading coefficients.

Remark 5.1.2. *Let $A\mathbf{x} = \mathbf{b}$ be a consistent system. Observe that*
$$\begin{aligned} &\textit{the number of pivots in } \text{rref}(A) \\ +\ &\textit{the number of free variables in the solution} \\ =\ &\textit{the number of columns of } A. \end{aligned}$$

Example 5.1.3.
$$\begin{cases} x_1 + \ \ x_2 + 2x_3 = 9 \\ 2x_1 + 4x_2 - 3x_3 = 1 \\ 3x_1 + 6x_2 - 5x_3 = 0 \end{cases}$$
Apply elementary row operations to $[A|\mathbf{b}]$ to obtain $[\text{rref}(A)|\mathbf{b}']$
$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right].$$
We solve $\text{rref}(A)\mathbf{x} = \mathbf{b}'$ to conclude that
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Example 5.1.4.
$$\begin{cases} x_1 - 2x_2 - \ \ x_3 + 3x_4 = 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 = 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 = 2 \end{cases}$$

Applying elementary row operations to $[A|\mathbf{b}]$ *to obtain*

$$[\text{rref}(A)|\mathbf{b}'] = \left[\begin{array}{cccc|c} 1 & -2 & 0 & 10/3 & 1 \\ 0 & 0 & 1 & 1/3 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{array}\right].$$

From the last row, we conclude that this system is inconsistent.

Example 5.1.5.

$$\begin{cases} \quad\quad 3x_2 - \ \ 6x_3 + 6x_4 + 4x_5 = -5 \\ 3x_1 - 7x_2 + \ \ 8x_3 - 5x_4 + 8x_5 = 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{cases}$$

Applying elementary row operations to $[A|\mathbf{b}]$ *we obtain*

$$[\text{rref}(A)|\mathbf{b}'] = \left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right].$$

Solving the system associated with this, we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_3 - 3x_4 - 24 \\ 2x_3 + 2x_4 - 7 \\ x_3(= s) \\ x_4(= t) \\ 4 \end{bmatrix} = s\begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix}.$$

Example 5.1.6.

$$\begin{cases} 2x_1 + \ \ 4x_2 - 2x_3 + 2x_4 + 4x_5 = 2 \\ \ \ x_1 + \ \ 2x_2 - \ \ x_3 + 2x_4 \qquad\quad = 4 \\ 3x_1 + \ \ 6x_2 - 2x_3 + \ \ x_4 + 9x_5 = 1 \\ 5x_1 + 10x_2 - 4x_3 + 5x_4 + 9x_5 = 9 \end{cases}$$

Applying elementary row operations to $[A|\mathbf{b}]$*, we obtain*

$$[\text{rref}(A)|\mathbf{b}'] = \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & 3 & 2 \\ 0 & 0 & 1 & 0 & -1 & 4 \\ 0 & 0 & 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Solve the system associated with this matrix to obtain

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_5 + 2 \\ x_2 (= s) \\ x_5 + 4 \\ 2x_5 + 3 \\ x_5 (= t) \end{bmatrix} = s \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 4 \\ 3 \\ 0 \end{bmatrix}.$$

5.2 $A\mathbf{x} = \mathbf{b}$ in terms of L_A

Let A be a $m \times n$ matrix over F and $\mathbf{b} \in F^m$. Recall that the map

$$L_A : F^n \to F^m, \quad L_A(\mathbf{x}) = A\mathbf{x}$$

is a linear transformation.

Theorem 5.2.1. *Let A be a $m \times n$ matrix over F. The following are equivalent.*

1. *A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent.*
2. *$\mathbf{b}$ is in the image of the map L_A*

$$\mathbf{b} \in \operatorname{im} L_A.$$

3. *$\mathbf{b}$ is in the subspace of F^m spanned by the columns A_j of A*

$$\mathbf{b} \in \operatorname{Span}(\{A_1, ..., A_n\}).$$

Proof. □

Corollary 5.2.2. *As a subspace of F^m,*

$$\operatorname{im} L_A = \operatorname{Span}(\{A_1, ..., A_n\}).$$

Corollary 5.2.2 also follows from Proposition 3.2.3, since the columns of A are the image of the elementary basis $\mathcal{E} = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$ of F^n under L_A.

$$L_A(\mathbf{e}_j) = A_j$$

For this reason, the image of the linear transformation L_A is often called the column space of A and denoted by $\operatorname{Col}(A)$.

5.3 Homogeneous systems

A system of linear equations $A\mathbf{x} = \mathbf{b}$ is called homogeneous, if $\mathbf{b} = \mathbf{0}$.

Proposition 5.3.1. *The solution set of a homogeneous equation $A\mathbf{x} = \mathbf{0}$ is exaclty the kernel of the linear transformation L_A.*

$$\ker L_A = \{\mathbf{x} \in F^n : A\mathbf{x} = \mathbf{0}\}.$$

For this reason, the kernel of the linear transformation L_A is often called the null space of A and denoted by $\mathrm{Null}(A)$.

Proposition 5.3.2. *Let $\mathbf{x}_p$ be a solution to a system $A\mathbf{x} = \mathbf{b}$. Every solution to $A\mathbf{x} = \mathbf{b}$ is of the form*

$$\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p$$

for some $\mathbf{x}_h \in \mathrm{Null}(A)$.

Proof. Let $\mathbf{x}$ be a general solution to $A\mathbf{x} = \mathbf{b}$. Then,

$$A(\mathbf{x} - \mathbf{x}_p) = A\mathbf{x} - A\mathbf{x}_p = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

and thus $\mathbf{x} - \mathbf{x}_p \in \mathrm{Null}(A)$. □

Example 5.3.3. *Find a basis of the solution space to each of the following homogeneous systems.*

1. $$\begin{cases} x_1 + x_2 + 2x_3 = 0 \\ 2x_1 + 4x_2 - 3x_3 = 0 \\ 3x_1 + 6x_2 - 5x_3 = 0 \end{cases}$$

2. $$\begin{cases} 3x_2 - 6x_3 + 6x_4 + 4x_5 = 0 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 = 0 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 0 \end{cases}$$

3. $$\begin{cases} 2x_1 + \ \ 4x_2 - 2x_3 + 2x_4 + 4x_5 = 0 \\ \ \ x_1 + \ \ 2x_2 - \ \ x_3 + 2x_4 \qquad\quad = 0 \\ 3x_1 + \ \ 6x_2 - 2x_3 + \ \ x_4 + 9x_5 = 0 \\ 5x_1 + 10x_2 - 4x_3 + 5x_4 + 9x_5 = 0 \end{cases}$$

Remark 5.3.4. *Let us define the rank and nullity of a matrix A by the rank and nullity of the linear transformation L_A, i.e., the dimensions of* $\operatorname{im} L_A$ *and* $\ker L_A$ *respectively.*

1. *Observe that in the solution to $A\mathbf{x} = \mathbf{0}$, the vectors multiplied by free variables form a basis of the solution space. Thus,*

$$\begin{aligned} \text{the nullity of } A &= \dim \ker L_A = \dim \operatorname{Null}(A) \\ &= \text{the number of free variables in the solution to } A\mathbf{x} = \mathbf{0}. \end{aligned}$$

2. *Then using the observation in Remark 5.1.2 with Theorem 3.3.1 applied to L_A,*

$$\begin{aligned} \text{the rank of } A &= \dim \operatorname{im} L_A = \dim \operatorname{Col}(A) \\ &= \text{the number of pivots in } \operatorname{rref}(A). \end{aligned}$$

The rank of a matrix A is often defined by the number of pivots in $\operatorname{rref}(A)$.

Chapter 6

Invertible matrix

가역행렬의 기본 성질을 공부하고 이를 선형변환과 연립 일차방정식 맥락에서 이해해봅시다. 정사각 행렬의 특성을 알려주는 값으로 행렬식을 정의하고 그 성질을 공부합시다.

6.1 Matrix inverse

Recall that a $n \times n$ matrix A is invertible, if there exists a $n \times n$ matrix X such that

$$XA = I_n \text{ and } AX = I_n.$$

We write A^{-1} for X and call it the inverse of A (Definition 4.2.5).

Proposition 6.1.1. *Let A be a $n \times n$ matrix.*

1. *If A is invertible then its inverse is unique.*
2. *If A is invertible and $XA = I_n$ then $AX = I_n$. Therefore, $X = A^{-1}$.*
3. *If A is invertible and $AX = I_n$ then $XA = I_n$. Therefore, $X = A^{-1}$.*

Proof. (1) Let X and X' be inverses of A. Then, by multiplying $AX = I_n$ by X',

$$X'(AX) = X'I_n$$
$$(X'A)X = X'.$$

Then, using $X'A = I_n$, we conclude that $X = X'$.

(2) By multiplying $XA = I_n$ by A,

$$A(XA) = AI_n$$
$$(AX)A = A.$$

Then, by multiplying A^{-1}, we obtain $((AX)A)A^{-1} = AA^{-1}$ and thus $AX = I_n$. □

Proposition 6.1.2. *Let A and B be $n \times n$ matrices.*

1. *If A is invertible then its inverse A^{-1} is invertible and*
$$(A^{-1})^{-1} = A.$$
2. *If A and B are invertible then their product AB is invertible and*
$$(AB)^{-1} = B^{-1}A^{-1}.$$
3. *If A is invertible then its transpose A^T is invertible and*
$$(A^T)^{-1} = (A^{-1})^T.$$

Proof. (1) From the definition of A^{-1},

$$A^{-1}A = AA^{-1} = I_n$$

and it shows that A is the inverse of A^{-1}. (2) Note that

$$(B^{-1}A^{-1})(AB) = (AB)(B^{-1}A^{-1}) = I_n.$$

(3) Using $(XY)^T = Y^T X^T$,

$$(A^{-1})^T A^T = (AA^{-1})^T = I_n, \quad \text{and} \quad A^T(A^{-1})^T = (A^{-1}A)^T = I_n.$$

□

Proposition 6.1.3. *A $n \times n$ matrix A is invertible if and only if the linear transformation*

$$L_A : F^n \to F^n, \quad \mathbf{x} \mapsto A\mathbf{x}$$

is bijective.

Proof. If a $n \times n$ matrix A is invertible then it is straightfoward to verify that the map defined by

$$L_{A^{-1}} : F^n \to F^n, \quad \mathbf{x} \mapsto A^{-1}\mathbf{x}$$

is the inverse map of L_A. Therefore, L_A is bijective.

Conversely, assume that the linear transformation L_A is bijective and L_A^{-1} is its inverse. Then L_A^{-1} is a linear transformation from F^n to F^n, and thus there is a matrix B such that

$$L_A^{-1} : F^n \to F^n, \quad \mathbf{x} \mapsto B\mathbf{x}.$$

Then, the compositions $L_A^{-1} \circ L_A$ and $L_A \circ L_A^{-1}$ are the identity map on F^n, and

$$(BA)\mathbf{x} = \mathbf{x} \text{ and } (AB)\mathbf{x} = \mathbf{x}$$

for all $\mathbf{x} \in F^n$, which implies that $BA = I_n$ and $AB = I_n$. Therefore, the matrix A is invertible. □

Theorem 6.1.4. *Let A be a $n \times n$ matrix. The following are equivalent.*

1. *A is invertible.*
2. *The map $L_A : F^n \to F^n$ is bijective.*
3. *The nullity of L_A is 0.*
4. *The rank of L_A is n.*
5. *The number of pivots in* $\operatorname{rref}(A)$ *is n.*
6. $\operatorname{rref}(A) = I_n$.
7. *The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only trivial solution.*
8. *The columns of A are linearly independent.*
9. *For every $\mathbf{b} \in F^n$, the system $A\mathbf{x} = \mathbf{b}$ is consistent.*
10. *For every $\mathbf{b} \in F^n$, $\mathbf{b} \in \operatorname{Col}(A)$.*

Proof. $(1) \Leftrightarrow (2)$ by Proposition 6.1.3. Note that (3) holds if and only if L_A is injective and (4) holds if and only if L_A is surjective. Therefore, (2) holds if and only if both (3) and (4) hold. We have $(3) \Leftrightarrow (4)$ by the rank-nullity theorem. Then use Remark 5.3.4 to verify that

$$(4) \Leftrightarrow (5) \Leftrightarrow (6), \quad (3) \Leftrightarrow (7) \Leftrightarrow (8), \quad \text{and} \quad (4) \Leftrightarrow (9) \Leftrightarrow (10).$$

□

Theorem 6.1.4는 행렬 A가 가역invertible 행렬에 대해 크게 세가지 관점에서 기술하고 있습니다.

a) 선형변환 L_A라는 함수의 관점에서,

b) 일차 연립 방정식 $A\mathbf{x} = \mathbf{b}$ 관점에서,

c) 행렬 A를 구성하는 n개의 종벡터들의 관점에서.

다음 장에서 배우는 행렬식을 이용하면, 행렬 A가 n^2개 숫자들의 나열이라는 관점에서, A의 가역성을 A의 원소 $(A)_{ij}$ 들을 이용하여 판정할 수 있습니다 (Theorem 6.4.2).

6.2 Elementary matrices and A^{-1}

6.2.1 Elementary matrices

Elementary matrices are $n \times n$ matrices obtained by applying elementary row operations to the identity matrix I_n. Note that the columns of I_n form the elementary basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ of F^n and the ith row of I_n is the transpose of $\mathbf{e}_i^T$ of the ith column $\mathbf{e}_i$ of I_n.

1. Exchange R_i and R_j in the identity matrix I_n:
$$I_n = \begin{bmatrix} \\ - & \mathbf{e}_i^T & - \\ \\ - & \mathbf{e}_j^T & - \\ \\ \end{bmatrix} \rightsquigarrow E = \begin{bmatrix} \\ - & \mathbf{e}_j^T & - \\ \\ - & \mathbf{e}_i^T & - \\ \\ \end{bmatrix}$$
2. Replace R_j with kR_j in the identity matrix I_n:
$$I_n = \begin{bmatrix} \\ - & \mathbf{e}_i^T & - \\ \\ \end{bmatrix} \rightsquigarrow E = \begin{bmatrix} \\ - & k\mathbf{e}_i^T & - \\ \\ \end{bmatrix}$$
3. Replace R_j with $kR_i + R_j$ in the identity matrix I_n:
$$I_n = \begin{bmatrix} \\ - & \mathbf{e}_i^T & - \\ \\ - & \mathbf{e}_j^T & - \\ \\ \end{bmatrix} \rightsquigarrow E = \begin{bmatrix} \\ - & \mathbf{e}_i^T & - \\ \\ - & k\mathbf{e}_i^T + \mathbf{e}_j^T & - \\ \\ \end{bmatrix}$$

If M' is a matrix obtained by applying an elementary row operation to a $n \times \ell$ matrix M then
$$M' = EM$$
where E is the elementary matrix obatined by applying the corresponding operation to I_n. Then, rref(A) can be obtained by multiplying A by elementary matrices (corresponding to elementary row operations) from the left hand side,
$$(E_k(\cdots(E_2(E_1A))\cdots)) = \text{rref}(A)$$
$$(E_k \cdots E_2E_1)A = \text{rref}(A) \tag{6.2.1}$$

6.2.2 Computing A^{-1}

Let P and Q be $n \times n$ matrices. Applying an elementary row operation to the augmented matrix $M = [P \mid Q]$ is equivalent to multiplying the corresponding elementary matrix E to $M = [P \mid Q]$. Then, it is

$$E[P \mid Q] = [EP \mid EQ].$$

Now, recall that a $n \times n$ matrix A is invertible if and only if $\mathrm{rref}(A) = I_n$, which is equivalent to say that there are elementary matrices $E_1, E_2, ..., E_k$ such that

$$(E_k \cdots E_2 E_1)A = I_n.$$

Multiplying them to the augmented matrix $[A \mid I_n]$, we obtain

$$\begin{aligned} &(E_k \cdots E_2 E_1)\,[A \mid I_n] \\ =&[(E_k \cdots E_2 E_1)A \mid (E_k \cdots E_2 E_1)I_n] \\ =&[I_n \mid (E_k \cdots E_2 E_1)] \\ =&[I_n \mid B]. \end{aligned}$$

Knowing that A is invertible and we have B such that $BA = I_n$, by Proposition 6.1.1 (or Proposition 6.2.3), we conclude that $B = A^{-1}$.

Proposition 6.2.1. *A $n \times n$ matrix A is invertible if and only if we can obtain $[I_n \mid B]$ by applying elementary row operations to $[A \mid I_n]$. In this case,*

$$B = A^{-1}.$$

6.2.3 LU decomposition

In (6.2.1), we observed that there are elementary matrices E_j such that

$$(E_k \cdots E_2 E_1)A = \mathrm{rref}(A).$$

Note that all these elementary matrices are invertible.

Definition 6.2.2. *An LU decomposition of a square matrix A is the product of a lower triangular matrix and an upper triangular matrix that is equal to A*

$$A = LU.$$

6.2.4 Left and right inverses

In Proposition 6.1.1, we showed that, knowing that A is invertible, a left inverse of A is indeed the inverse of A. Now, for later use, we show that the statement still holds without the invertibility condition.

Proposition 6.2.3. *Let A be a $n \times n$ matrix.*

1. *There is a $n \times n$ matrix N such that $NA = I_n$ if and only if $A\mathbf{x} = \mathbf{b}$ has a solution for all $\mathbf{b} \in F^n$.*
2. *A left inverse of a A is a right inverse of A. That is, if there is a $n \times n$ matrix N such that $NA = I_n$ then $AN = I_n$. Thus, $N = A^{-1}$.*
3. *A right inverse of A is a left inverse of A. That is, if there is a $n \times n$ matrix M such that $AM = I_n$ then $MA = I_n$. Thus $M = A^{-1}$.*

Proof. (1) ($\Rightarrow$) For all $\mathbf{b} \in F^n$, $\mathbf{x} = N\mathbf{b}$ is a solution to $A\mathbf{x} = \mathbf{b}$. ($\Leftarrow$) Suppose that there is no N satisfying $NA = I_n$. Then, because rref(A) can be obtained by multiplying A by elementary matrices (corresponding to elementary row operations) from the left hand side,

$$(E_k(\cdots(E_2(E_1A))\cdots)) = \text{rref}(A)$$
$$(E_1 \cdots E_2E_1)A = \text{rref}(A)$$

we have rref$(A) \neq I_n$. In this case, rref(A) contains a zero row and thus we can find $\mathbf{b} \in F^n$ such that $A\mathbf{x} = \mathbf{b}$ is inconsistent.

(2) Suppose that there is N such that $NA = I_n$. Then, from statement (1), for any $\mathbf{b} \in F^n$, there is $\mathbf{x} \in F^n$ such that $A\mathbf{x} = \mathbf{b}$. By multiplying $A\mathbf{x} = \mathbf{b}$ by N,

$$N(A\mathbf{x}) = N\mathbf{b}$$
$$(NA)\mathbf{x} = N\mathbf{b},$$

and thus $\mathbf{x} = N\mathbf{b}$. This shows that

$$\mathbf{b} = A\mathbf{x} = A(N\mathbf{b}) = (AN)\mathbf{b}.$$

Writing N' for the product AN, we have $\mathbf{b} = N'\mathbf{b}$ for all $\mathbf{b} \in F^n$. This shows that $N' = I_n$ and thus $AN = I_n$. From $NA = AN = I_n$, we conclude that A is invertible and $N = A^{-1}$.

(3) If $AM = I_n$ then $M^T A^T = I_n$. Thus, by statement (2), A^T is invertible and its inverse is M^T. Then, by Proposition 6.1.2, $A = (A^T)^T$ is invertible and its inverse is $M = (M^T)^T$. □

6.3 Determinant

The determinant is a function defined on the set of square matrices

$$\det : M_n(F) \to F, \quad A \mapsto \det(A).$$

It can be considered a function of the rows of a matrix:

$$\det(A) = \det \begin{bmatrix} - & R_1 & - \\ & \vdots & \\ - & R_n & - \end{bmatrix}$$

where $R_k = R_k(A)$ is the kth row of A. Similarly, the determinant can be considered a function of columns of a matrix.

Let us begin with the determinant of 1×1 and 2×2 matrices:

$$\det \begin{bmatrix} a \end{bmatrix} = a \quad \text{and} \quad \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

Definition 6.3.1 (Row expansion). *We define the determinant* $\det(A)$ *of a square matrix* $A = (a_{ij}) \in M_n(F)$ *inductively as follows.*

1. *For* $A = (a) \in M_1(F)$, $\det(A) = a$.
2. *Let* $n \geq 2$. *For each* $1 \leq i \leq n$,
$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{ij})$$
where A^{ij} *is the* $(n-1) \times (n-1)$ *matrix obtained by erasing the* i*th row and* j*th column of* A.

It turns out that the formula of $\det(A)$ provides the same value for all i. In Proposition 6.3.4, we will give a formula which does not involve i.

Lemma 6.3.2. 1. *For* $n \geq 1$, *the determinant of the identity matrix* I_n *is one.*
$$\det(I_n) = 1.$$

2. *If any two rows in* A *are exchanged then the sign of the determinant*

changes. That is, as a function of the rows of a matrix, the determinant is alternating.

$$\det \begin{bmatrix} & \vdots & \\ - & R_j & - \\ & \vdots & \\ - & R_i & - \\ & \vdots & \end{bmatrix} = -\det \begin{bmatrix} & \vdots & \\ - & R_i & - \\ & \vdots & \\ - & R_j & - \\ & \vdots & \end{bmatrix}$$

3. *As a function of the rows of a matrix, the determinant is multilinear, that is,*

$$\det \begin{bmatrix} & \vdots & \\ - & R_i + R_i' & - \\ & \vdots & \end{bmatrix} = \det \begin{bmatrix} & \vdots & \\ - & R_i & - \\ & \vdots & \end{bmatrix} + \det \begin{bmatrix} & \vdots & \\ - & R_i' & - \\ & \vdots & \end{bmatrix}$$

$$\det \begin{bmatrix} & \vdots & \\ - & \alpha R_i & - \\ & \vdots & \end{bmatrix} = \alpha \det \begin{bmatrix} & \vdots & \\ - & R_i & - \\ & \vdots & \end{bmatrix}$$

for all R_i, R_i' and $\alpha \in F$.

Proof. Use mathematical inductionon n. See, for example, Theorem 4.3 in Friedberg (5th edition). □

Let $\mathcal{E} = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$ be the elementary basis of F^n. Using Lemma 6.3.2, we can show that as a function of the rows of a matrix

$$\det \begin{bmatrix} - & \mathbf{e}_{i_1}^T & - \\ & \vdots & \\ - & \mathbf{e}_{i_n}^T & - \end{bmatrix} = \begin{cases} \pm 1 & \text{if } \{i_1, ..., i_n\} = \{1, ..., n\} \\ 0 & \text{otherwise.} \end{cases}$$

Then, for a permutation $\sigma \in \mathfrak{S}_n$,

$$\det \begin{bmatrix} - & \mathbf{e}_{\sigma(1)}^T & - \\ & \vdots & \\ - & \mathbf{e}_{\sigma(n)}^T & - \end{bmatrix} = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

Here, σ is even or odd, if we need even or odd number of transpositions (exchanges of two elements) respectively to obtain $(1, 2, ..., n)$ from $(\sigma(1), ..., \sigma(n))$. It is known that this parity is well-defined regardless of actual transpositions we use. Now, let us define the sign (or the signature) of σ by

$$sgn(\sigma) = \det \begin{bmatrix} - & \mathbf{e}_{\sigma(1)}^T & - \\ & \vdots & \\ - & \mathbf{e}_{\sigma(n)}^T & - \end{bmatrix}.$$

Then, it is immediate to see that $sgn(\sigma^{-1}) = sgn(\sigma)$.

Lemma 6.3.3. *Let $R_k = R_k(A)$ be the kth row of a matrix A. For any permutation $\sigma \in \mathfrak{S}_n$,*

$$\det \begin{bmatrix} - & R_{\sigma(1)} & - \\ - & R_{\sigma(2)} & - \\ & \vdots & \\ - & R_{\sigma(n)} & - \end{bmatrix} = sgn(\sigma) \det \begin{bmatrix} - & R_1 & - \\ - & R_2 & - \\ & \vdots & \\ - & R_n & - \end{bmatrix}.$$

We remark that there is another way of computing $sgn(\sigma)$: let us call a pair (i, j) the inversion of σ, if $i < j$ and $\sigma(i) > \sigma(j)$. Writing $inv(\sigma)$ for the number of inversions of σ, it is known that

$$sgn(\sigma) = (-1)^{inv(\sigma)}.$$

Proposition 6.3.4. *The determinant of $A = (a_{ij}) \in M_n(F)$ is*

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma)\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Proof. The ith row of A is $R_i = \sum_{j=1}^n a_{i,j}\mathbf{e}_j^T$, Then using the mutilinearity of the determinant map shown in Lemma 6.3.2,

$$\begin{aligned}
\det(R_1, ..., R_n) &= \det\left(\sum_{j_1=1}^{n} a_{1,j_1}\mathbf{e}_{j_1}^T, ..., \sum_{j_n=1}^{n} a_{n,j_n}\mathbf{e}_{j_n}^T\right) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} \det\left(a_{1,j_1}\mathbf{e}_{j_1}^T, ..., a_{n,j_n}\mathbf{e}_{j_n}^T\right) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1,j_1} \cdots a_{n,j_n} \det\left(\mathbf{e}_{j_1}^T, ..., \mathbf{e}_{j_n}^T\right) \\
&= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) a_{1,j_1} \cdots a_{n,j_n} \qquad \text{where } \sigma(k) = j_k \text{ for all } k.
\end{aligned}$$

□

Theorem 6.3.5. *For every square matrix $A = (a_{ij}) \in M_n(F)$,*

$$\det(A^T) = \det(A).$$

Proof. Let $A^T = (b_{ij})$ thus $b_{ij} = a_{ji}$.

$$\begin{aligned}
\det(A^T) &= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} \\
&= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) a_{\sigma(1),1} \cdots b_{\sigma(n),n} \\
&= \sum_{\sigma^{-1} \in \mathfrak{S}_n} sgn(\sigma^{-1}) a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \\
&= \det(A).
\end{aligned}$$

□

Therefore, we can think of the determinant is a function of the columns of a matrix

$$\det(A) = \det\begin{bmatrix} | & & | \\ C_1 & \cdots & C_n \\ | & & | \end{bmatrix}$$

where $C_k = C_k(A)$ is the kth column of A.

Corollary 6.3.6. 1. *(Column expansion). For each* $1 \le j \le n$,

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{ij}).$$

2. *As a function of the columns of a matrix, the determinant is alternating.*

3. *As a function of the column of a matrix, the determinant is multilinear.*

4. *For any permutation* $\sigma \in \mathfrak{S}_n$,

$$\det \begin{bmatrix} | & | & & | \\ C_{\sigma(1)} & C_{\sigma(2)} & \cdots & C_{\sigma(n)} \\ | & | & & | \end{bmatrix} = sgn(\sigma) \det \begin{bmatrix} | & | & & | \\ C_1 & C_2 & \cdots & C_n \\ | & | & & | \end{bmatrix}.$$

Recall that the trace map is compatible with matrix addition (Example 3.1.2). Now we show that the determinant map is compatible with matrix multiplication.

Theorem 6.3.7. *For* $A, B \in M_n(F)$,

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

Proof. For a matrix M, write M_j for the jth column of M. The jth column of the product AB is

$$(AB)_j = \begin{bmatrix} \sum_k a_{1k} b_{kj} \\ \sum_k a_{2k} b_{kj} \\ \vdots \\ \sum_k a_{nk} b_{kj} \end{bmatrix} = b_{1j} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + b_{2j} \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + b_{nj} \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = \sum_{i=1}^{n} b_{ij} A_i$$

See also (4.2.2). Now considering the determinant as a function of columns of a matrix,

$$\begin{aligned}
\det(AB) &= \det\left(\sum_{i_1=1}^{n} b_{i_1,1}A_{i_1}, ..., \sum_{i_n=1}^{n} b_{i_n,n}A_{i_n}\right) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1,1}\cdots b_{i_n,n} \det(A_{i_1}, ..., A_{i_n}) \\
&= \sum_{\sigma\in\mathfrak{S}_n} b_{\sigma(1),1}\cdots b_{\sigma(n),n} sgn(\sigma) \det(A_1, ..., A_n) \\
&= \det(A) \sum_{\sigma\in\mathfrak{S}_n} sgn(\sigma) b_{\sigma(1),1}\cdots b_{\sigma(n),n} \\
&= \det(A)\det(B).
\end{aligned}$$

□

Corollary 6.3.8. *Let $A \in M_n(F)$.*

1. *If A is invertible then $\det(A) \neq 0$ and*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

2. $\det(A) = 0$ *if and only if* $\det(\mathrm{rref}(A)) = 0$.

Proof. (1) By applying Theorem 6.3.7 to $AA^{-1} = I_n$,

$$\det(AA^{-1}) = \det(A)\det(A^{-1}) = 1.$$

(2) By Theorem 6.3.7, if E_j are elementary matrices such that

$$E_k \cdots E_2 E_1 A = \mathrm{rref}(A)$$

then

$$\det(E_k)\cdots\det(E_1)\det(E_1)\det(A) = \det(\mathrm{rref}(A)).$$

By noting that $\det(E_j) \neq 0$, we have $\det(A) = 0$ if and only if $\det(\mathrm{rref}(A)) = 0$. □

The number of pivots in $\mathrm{rref}(A)$ is n if and only if $\mathrm{rref}(A) = I_n$, in which case $\det(\mathrm{rref}(A)) = 1$ and thus $\det(A) \neq 0$. If the number of pivots in $\mathrm{rref}(A)$ is strictly less than n then $\mathrm{rref}(A)$ contains a zero row and therefore $\det(\mathrm{rref}(A)) = 0$. This discussion combined with Theorem 6.1.4 (1) ⇔ (6) shows that the converse of Corollary 6.3.8 (1) is also true. We will give a more constructive proof in Theorem 6.4.2.

Proposition 6.3.9. *If two square matrices A and B are similar, i.e., there is an invertible matrix Q such that $B = Q^{-1}AQ$ then*

$$tr(A) = tr(B) \quad \textit{and} \quad \det(A) = \det(B).$$

Proof. First we note that $tr(XY) = tr(YX)$ for any $n \times n$ matrices X and Y.

$$tr(XY) = \sum_{i=1}^{n}(XY)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n}(X)_{ik}(Y)_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n}(Y)_{ki}(X)_{ik} = tr(YX).$$

Then, $tr(B) = tr(Q^{-1}AQ) = tr(AQQ^{-1}) = tr(A)$. Next, using Theorem 6.3.7,

$$\begin{aligned}\det(B) = \det(Q^{-1}AQ) &= \det(Q^{-1})\det(A)\det(Q)\\ &= \det(Q)^{-1}\det(A)\det(Q) = \det(A).\end{aligned}$$

□

We define the trace and determinant of a linear transformation $T : V \to V$ by the trace and determinant of the matrix $[T]_{\mathcal{B}}$ with relative to any basis $\mathcal{B}$. By Proposition 6.3.9, they are well defined regardless of $\mathcal{B}$.

The determinant a 2×2 real matrix $A = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix}$ gives the “signed” area of a paralleogram formed by two vectors $\mathbf{v}_1 = \begin{bmatrix}a & c\end{bmatrix}^T$ and $\mathbf{v}_2 = \begin{bmatrix}b & d\end{bmatrix}^T$ in $\mathbb{R}^2$:

$$\det\begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

The determinant a 3×3 real matrix $A = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix}$ gives the “signed” volumn of a parallelepiped formed by three vectors $\mathbf{v}_i = \begin{bmatrix}a_i & b_i & c_i\end{bmatrix}^T$ in $\mathbb{R}^3$:

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

Note that it is the box product (or scalar triple product) given in Calculus.

Considering the corresponding linear transformation

$$L_A : \mathbb{R}^n \to \mathbb{R}^n, \qquad \mathbf{x} \mapsto A\mathbf{x}$$

the determinant of A

$$\det(A) = \det \begin{bmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{bmatrix} \quad \text{where } A_j = L_A(\mathbf{e}_j)$$

measures how much the map L_A deforms the unit cube formed by $\mathbf{e}_1, ..., \mathbf{e}_n$.

6.4 Cramer's rule and A^{-1}

Let us consider a system of n linear equations for n unknowns

$$A\mathbf{x} = \mathbf{b}.$$

Theorem 6.4.1. *Let A be a $n \times n$ matrix with nonzero determinant. Then, for any $\mathbf{b} \in F^n$, the system of linear equations $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}$ whose ith entry is*

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det(A)}$$

where $A_i(\mathbf{b})$ is the matrix obtained by replacing the ith column of A with $\mathbf{b}$.

Proof. Let B be the matrix obtained from I_n by replacing ith column by $\mathbf{x}$. Note that $A\mathbf{e}_i$ is the ith column of A. Then, the product AB is

$$A \begin{bmatrix} | & & | & & | \\ \mathbf{e}_1 & \cdots & \mathbf{x} & \cdots & \mathbf{e}_n \\ | & & | & & | \end{bmatrix} = \begin{bmatrix} | & & | & & | \\ A_1 & \cdots & A\mathbf{x} & \cdots & A_n \\ | & & | & & | \end{bmatrix}.$$

Then, after replacing $A\mathbf{x}$ by $\mathbf{b}$, take the determinant to obtain

$$\det(A)\det(B) = \det(A_i(\mathbf{b})).$$

The statement follows from $\det(B) = x_i$. □

Theorem 6.4.2. *A square matrix A is invertible if and only if $\det(A) \neq 0$.*

Proof. ($\Rightarrow$) This is the first statement of Corollary 6.3.8. ($\Leftarrow$) Let A be a $n \times n$ matrix with nonzero determinant. By Theorem 6.4.1, the systems $A\mathbf{x} = \mathbf{e}_j$ have unique solutions

$$A\mathbf{x}_1 = \mathbf{e}_1,\ A\mathbf{x}_2 = \mathbf{e}_2,\ \dots,\ A\mathbf{x}_n = \mathbf{e}_n.$$

Then, we have

$$A \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix} = I_n.$$

Therefore, by Proposition 6.2.3, the inverse of A is

$$A^{-1} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix}.$$

□

Theorem 6.4.3. *Let A be a $n \times n$ matrix. If $\det(A) \neq 0$ then the (i, j) entry of A^{-1} is*

$$\frac{(-1)^{i+j} \det(A^{ji})}{\det(A)}$$

where A^{ji} is the $(n-1) \times (n-1)$ matrix obtained from A by erasing the jth row and the ith column.

Proof. By solving $A\mathbf{x}_j = \mathbf{e}_j$ for $1 \leq j \leq n$, we can find a $n \times n$ matrix X such that $AX = I_n$. Then, the ith entry of the vector $\mathbf{x}_j$ is, by Cramer's rule,

$$x_{ij} = \frac{\det A_i(\mathbf{e}_j)}{\det(A)}.$$

The statement follows from the observation

$$\det A_i(\mathbf{e}_j) = (-1)^{i+j} \det(A^{ji}).$$

□

Example 6.4.4.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1} = \frac{1}{\det(A)} \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$