LINEAR ALGEBRA II

Hiroshi SUZUKI*
Division of Natural Sciences
International Christian University

February 20, 2007

1 Euclidean n-Space

Definition 1.1 If n is a positive integer, then an ordered n-tuple is a sequence of n real numbers $(a_1, a_2, \ldots, a_n)$. The set of all ordered n-tuples is called n-space and is denoted by $\boldsymbol{R}^n$.

Definition 1.2 Two vectors $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ in $\boldsymbol{R}^n$ are called *equal* if

$$u_1 = v_2, u_2 = v_2, \ldots, u_n = v_n.$$

The *sum* $\boldsymbol{u} + \boldsymbol{v}$ is defined by

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$

and if k is any scalar, the *scalar multiple* $k\boldsymbol{u}$ is defined by

$$k\boldsymbol{u} = (ku_1, ku_2, \ldots, ku_n).$$

Let $\boldsymbol{0} = (0, 0, \ldots, 0) \in \boldsymbol{R}^n$, $-\boldsymbol{u} = (-u_1, -u_2, \ldots, -u_n)$ and $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{v} + (-\boldsymbol{u})$ or, in terms of components,

$$\boldsymbol{v} - \boldsymbol{u} = (v_1 - u_1, v_2 - u_2, \ldots, v_n - u_n).$$

Theorem 1.1 (4.1.1) *Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ be vectors in $\boldsymbol{R}^n$ and k and m scalars. Then:*

(a) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$. *(Commutativity)*

(b) $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ *(Associativity)*

(c) $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{u} = \boldsymbol{u}$

(d) $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$*; that is* $\boldsymbol{u} - \boldsymbol{u} = \boldsymbol{0}$.

(e) $k(m\boldsymbol{u}) = (km)\boldsymbol{u}$.

(f) $k(\boldsymbol{u} + \boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$.

(g) $(k + m)\boldsymbol{u} = k\boldsymbol{u} + m\boldsymbol{u}$.

(h) $1\boldsymbol{u} = \boldsymbol{u}$.

Definition 1.3 Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ be vectors in $\boldsymbol{R}^n$. Then *the Euclidean Inner Product* $\boldsymbol{u} \cdot \boldsymbol{v}$ is defined by

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n,$$

the Euclidean norm (or Euclidean length) of a vector $\boldsymbol{u}$ is defined by

$$\|\boldsymbol{u}\| = (\boldsymbol{u} \cdot \boldsymbol{u})^{1/2} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2},$$

and the Euclidean distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined by

$$\begin{aligned} d(\boldsymbol{u}, \boldsymbol{v}) &= \|\boldsymbol{u} - \boldsymbol{v}\| \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}. \end{aligned}$$

Theorem 1.2 (4.1.2) *Let $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ be vectors in $\boldsymbol{R}^n$ and k a scalar. Then:*

(a) $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$.

(b) $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$.

(c) $(k\boldsymbol{u}) \cdot \boldsymbol{v} = k(\boldsymbol{u} \cdot \boldsymbol{v})$.

(d) $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$. *Further,* $\boldsymbol{v} \cdot \boldsymbol{v} = 0$ *if and only if* $\boldsymbol{v} = \boldsymbol{0}$.

Cauchy-Schwarz Inequality in $\boldsymbol{R}^n$.

Theorem 1.3 (4.1.3) *Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ be vectors in $\boldsymbol{R}^n$. Then*

$$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\| \|\boldsymbol{v}\|$$

Equality holds if and only if $\boldsymbol{v} = k\boldsymbol{u}$ for some real k or $\boldsymbol{u} = \boldsymbol{0}$.

Theorem 1.4 (4.1.4) *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in $\boldsymbol{R}^n$ and k a scalar. Then:*

(a) $\|u\| \geq 0$.

(b) $\|\boldsymbol{u}\| = 0$ *if and only if* $\boldsymbol{u} = \boldsymbol{0}$.

(c) $\|k\boldsymbol{u}\| = |k| \|\boldsymbol{u}\|$.

(d) $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. *(Triangle inequality)*

*E-mail:hsuzuki@icu.ac.jp

Theorem 1.5 (4.1.4) *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in $\boldsymbol{R}^n$ and k a scalar. Then:*

(a) $d(\boldsymbol{u}, \boldsymbol{v}) \geq 0$.

(b) $d(\boldsymbol{u}, \boldsymbol{v}) = 0$ *if and only if* $\boldsymbol{u} = \boldsymbol{v}$.

(c) $d(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{v}, \boldsymbol{u})$.

(d) $d(\boldsymbol{u}, \boldsymbol{v}) \leq d(\boldsymbol{u}, \boldsymbol{w}) + d(\boldsymbol{w}, \boldsymbol{v})$.
(Triangle inequality)

2 Linear Transformations from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$

Definition 2.1 Let X and Y be sets. A *function* (or *mapping*) f is a rule that associates with each element $a \in X$ one and only element $b \in Y$.

- $f : X \to Y$ $(a \mapsto b = f(a))$.
- b is the *image* of a under f, or $f(a)$ is the value of f at a.
- X is the *domain* of f and Y is the *codomain* of f.
- $\mathrm{Im} f = \{f(a) \mid a \in X\}$ is called the *range* of f.

Two functions (mappings) $f_1 : X_1 \to Y_1$ and $f_2 : X_2 \to Y_2$ are equal if $X_1 = X_2$, $Y_1 = Y_2$ and $f_1(a) = f_2(a)$ for all $a \in X_1 = X_2$.

Definition 2.2 If the domain of a function T is $\boldsymbol{R}^n$ and the codomain R^m then T is called a transformation from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$.

A mapping $T : \boldsymbol{R}^n \to \boldsymbol{R}^m$ is called a *linear transformation* if

$$T : \boldsymbol{R}^n \to \boldsymbol{R}^m \; (\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mapsto T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right)$$

$$= \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{bmatrix})$$

Let

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \; A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ & & \cdots & \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

Then the linear transformation can be written as

$$T : \boldsymbol{R}^n \to \boldsymbol{R}^m \; (\boldsymbol{x} \mapsto A\boldsymbol{x}).$$

The matrix $A = [a_{i,j}]$ is called the *standard matrix* of T and write $A = [T]$.

Conversely if A is an $m \times n$ matrix and the mapping from R^n to R^m is defined by $\boldsymbol{x} \mapsto A\boldsymbol{x}$, then the linear transformation is denoted by T_A. In particular $[T_A] = A$.

Theorem 2.1 *Let $T_1 : \boldsymbol{R}^n \to \boldsymbol{R}^m$ and $T_2 : \boldsymbol{R}^m \to \boldsymbol{R}^\ell$ be linear transformations. Then the composition of T_2 with T_1 defined by*

$$T_2 \circ T_1 : \boldsymbol{R}^n \to \boldsymbol{R}^\ell \; (\boldsymbol{x} \mapsto T_2(T_1(\boldsymbol{x}))).$$

is a linear transformation and $[T_2 \circ T_1] = [T_2][T_1]$.

Theorem 2.2 (4.3.2) *A transformation $T : \boldsymbol{R}^n \to \boldsymbol{R}^m$ is linear if and only if the following hold for all $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{R}^n$ and for every scalar c.*
(a) $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ (b) $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$.

Corollary 2.3 (4.3.3) *If T is a linear transformation from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$ and $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, then*

$$[T] = [T\boldsymbol{e}_1, T\boldsymbol{e}_2, \ldots, T\boldsymbol{e}_n].$$

Definition 2.3 Let $f : X \to Y$ be a function (or mapping).

(a) If $\mathrm{Im}(f) = f(X) = Y$, then f is said to be surjective or onto.

(b) If $f(a) \neq f(a')$ whenever $a \neq a'$, f is said to be injective or one-to-one. f is injective iff $f(a) = f(a')$ implies $a = a'$ for all $a, a' \in X$.

(c) If f is one-to-one and onto, f is said to be bijective.

Theorem 2.4 (2.3.6) *If A is an $n \times n$ matrix, then the following statements are equivalent.*

(a) *A is invertible.*

(b) *$A\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution.*

(c) *The reduced echelon form of A is I_n.*

(d) *A can be expressed as a product of elementary matrices.*

(e) *$A\boldsymbol{x} = \boldsymbol{b}$ is consistent for every $n \times 1$ matrix $\boldsymbol{b}$.*

(f) *$A\boldsymbol{x} = \boldsymbol{b}$ has exactly one solution for every $n \times 1$ matrix $\boldsymbol{b}$.*

(g) $\det(A) \neq 0$.

Theorem 2.5 (4.3.1) *If A is an $n \times n$ matrix and $T_A : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is multiplication by A, then the following statements are equivalent.*

(a) *A is invertible.*

(b) *T_A is surjective.*

(c) *T_A is injective.*

(d) *T_A is bijective.*

3 Vector Spaces and Subspaces

3.1 Definition of Vector Spaces

In the following K denotes either the real number field $\boldsymbol{R}$, the set of real numbers with two binary operations, i.e., addition and multiplication, or the complex number field $\boldsymbol{C}$. K can be replaced by any algebraic structure called a *field* but assume $K = \boldsymbol{R}$ unless otherwise stated. Elements of K are called scalars.

$K = \{0, 1\}$ with addition and multiplication defined by $0+0 = 0$, $0+1 = 1+0 = 1$, $1+1 = 0$, and $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$, $1 \cdot 1 = 1$ is another example of a field.

Definition 3.1 [**Vector Space Axioms**] Let (K be a field and let) V be a set on which two operations are defined: additions and multiplication by scalars (numbers). (By *addition* we mean a rule for associating with each pair of elements $\boldsymbol{u}, \boldsymbol{v} \in V$ an element $\boldsymbol{u} + \boldsymbol{v} \in V$, called the *sum* of $\boldsymbol{u}$ and $\boldsymbol{v}$, by *scalar multiplication* we mean a rule for associating with each scalar k and each element $\boldsymbol{u} \in V$ an element $k\boldsymbol{u} \in V$, called the *scalar multiple* of $\boldsymbol{u}$ by k.) If the following axioms are satisfied, then we call V a *vector space* (over K) and we call the elements in V *vectors*.

1. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are elements in V, then $\boldsymbol{u} + \boldsymbol{v}$ is in V.
2. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$.
3. $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$.
4. There is an element $\boldsymbol{0} \in V$, called a *zero vector* for V, such that $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$ for all $\boldsymbol{u} \in V$.
5. For each $\boldsymbol{u} \in V$, there is an element $-\boldsymbol{u} \in V$, called a *negative* of $\boldsymbol{u}$, such that $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$.
6. If k is a scalar and $\boldsymbol{u}$ is an element in V, then $k\boldsymbol{u}$ is in V.
7. $k(\boldsymbol{u} + \boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$ and any scalar k.
8. $(k + m)\boldsymbol{u} = k\boldsymbol{u} + m\boldsymbol{u}$ for any vector $\boldsymbol{u} \in V$ and all scalars k and m.
9. $k(m\boldsymbol{u}) = (km)\boldsymbol{u}$ for any vector $\boldsymbol{u} \in V$ and all scalars k and m.
10. $1\boldsymbol{u} = \boldsymbol{u}$ for any vector $\boldsymbol{u} \in V$.

Vector spaces over $\boldsymbol{R}$ are called *real vector spaces* and vector spaces over $\boldsymbol{C}$ *complex vector spaces.*

Proposition 3.1 (5.1.1) *Let V be a vector space, $\boldsymbol{u}$ a vector in V, and k a scalar; then:*

(a) $0\boldsymbol{u} = \boldsymbol{0}$.

(b) $k\boldsymbol{0} = \boldsymbol{0}$.

(c) $(-1)\boldsymbol{u} = -\boldsymbol{u}$.

(d) *If $k\boldsymbol{u} = \boldsymbol{0}$, then $k = 0$ or $\boldsymbol{u} = \boldsymbol{0}$.*

3.2 Subspaces

Definition 3.2 A subset W of a vector space V is called a *subspace* of V if W is itself a vector space under the addition and scalar multiplication defined on V.

Theorem 3.2 (5.2.1) *If W is a nonempty subset of a vector space V, then W is a subspace of V if and only if the following conditions hold.*

(a) *$\boldsymbol{u} + \boldsymbol{v} \in W$ for all $\boldsymbol{u}, \boldsymbol{v} \in W$.*

(b) *$k\boldsymbol{u} \in W$ for all $\boldsymbol{u} \in W$ and all scalars k.*

Proposition 3.3 (5.2.2) *Let A be an $m \times n$ matrix, and $T = T_A$ a linear transformation defined by*

$$T : \boldsymbol{R}^n \to \boldsymbol{R}^m \ (\boldsymbol{x} \mapsto A\boldsymbol{x}).$$

Then $W = \{\boldsymbol{v} \in \boldsymbol{R}^n \mid T(\boldsymbol{v}) = \boldsymbol{0}\}$ is a subspace of a vector space $V = \boldsymbol{R}^n$. W is called the kernel of the linear transformation T and is denoted by $\mathrm{Ker}(T)$.

Definition 3.3 [**Linear Combination**] A vector $\boldsymbol{w}$ is called a *linear combination* of the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ if it can be expressed in the form

$$\boldsymbol{w} = k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_r\boldsymbol{v}_r$$

where $k_1, k_2, \ldots, k_r$ are scalars.

Theorem 3.4 (5.2.3) *If $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ are vectors in a vector space V, then*

(a) *The set W of all linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ is a subspace of V.*

(b) *W is the smallest subspace of V that contains $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ in the sense that every other subspace of V that contains $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ must contain W.*

Definition 3.4 If $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ is a set of vectors in a vector space V, then the subspace W of V consisting of all linear combinations of the vectors in S is called the *space spanned* by $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$, and we say that the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ *span* W. To indicate that W is the space spanned by the vectors in the set $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$, we write

$$W = \mathrm{Span}(S) \ \text{ or } \ W = \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}.$$

4 Linear Independence and Basis

4.1 Linear Independence

Definition 4.1 Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ be a nonempty set of vectors. If the equation

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_r\boldsymbol{v}_r = \boldsymbol{0}$$

has only one solution, namely, $k_1 = k_2 = \cdots = k_r = 0$, then S is called a *linearly independent* set. If there are other solutions, then S is called a *linearly dependent* set.

Proposition 4.1 (5.3.1, 5.4.1) *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ be a nonempty set of vectors. Then the following are equivalent.*

(a) *S is a linearly independent set.*

(b) *No vector in S is expressible as a linear combination of the other vectors in S.*

(c) *For each vector $\boldsymbol{v}$, $k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2+\cdots+k_r\boldsymbol{v}_r = \boldsymbol{v}$ has at most one solution, i.e., if*

$$k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2+\cdots+k_r\boldsymbol{v}_r = k'_1\boldsymbol{v}_1+k'_2\boldsymbol{v}_2+\cdots+k'_r\boldsymbol{v}_r$$

then $k_1 = k'_1$, $k_2 = k'_2, \ldots, k_r = k'_r$.

Theorem 4.2 (1.2.1) *A homogeneous system of linear equations with more unknowns than equations has infinitely many solutions.*

Theorem 4.3 (5.3.3) *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ be a set of vectors in $\boldsymbol{R}^n$. If $r > n$, then S is linearly dependent. In particular, if $A = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r]$, then a system of linear equation $A\boldsymbol{x} = \boldsymbol{0}$ has a nonzero solution.*

Proposition 4.4 (5.3.4) *If the functions $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$ have $n-1$ continuous derivatives on the interval (a, b), and if the Wronskian of these functions is not identically zero on (a, b), then these functions form a linearly independent vectors in $C^{(n-1)}(a, b)$.*

4.2 Basis and Dimension

Definition 4.2 If V is a vector space and $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a set of vectors fo V, then S is called a *basis* for V if the following two conditions hold:

(a) S is linearly independent.

(b) S spans V, i.e., every vector in V can be written as a linear combination of vectors in S.

V is called *finite-dimensional* if it contains a finite set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ that forms a basis. If no such set exists, V is called *infinite-dimensional.*

Theorem 4.5 (5.4.2) *Let V be a finite-dimensional vector space, and let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis.*

(a) *If a set has more than n vectors, then it is linearly dependent.*

(b) *If a set has fewer than n vectors, then it does not span V.*

Corollary 4.6 (5.4.3) *All bases for a finite-dimensional vector space have the same number of vectors.*

Definition 4.3 The *dimension* of a finite-dimensional vector space V, denoted by $\dim(V)$, is defined to be the number of vectors in a basis for V. (In addition, we define the zero vector space to have dimension zero.)

Proposition 4.7 (5.4.4) *Let S be a nonempty set of vectors in a vector space V.*

(a) *If S is a linearly independent set, and $\boldsymbol{v} \notin \mathrm{Span}(S)$, then $S \cup \{\boldsymbol{v}\}$ is a linearly independent set.*

(b) *If $\boldsymbol{v}$ is a vector in S that is expressible as a linear combination of other vectors in S, then $\mathrm{Span}(S \setminus \{\boldsymbol{v}\}) = \mathrm{Span}(S)$.*

Theorem 4.8 (5.4.5, 5.4.6, 5.4.7) *Let V be an n-dimensional vector space, and S a set of vectors in V*

(a) *Suppose S has exactly n vectors. Then S is linearly independent if and only if S spans V.*

(b) *If S spans V but not a basis for V, then S can be reduced to a basis for V by removing appropriate vectors from S.*

(c) *If S is linearly independent that is not already a basis for V, then S can be enlarged to a basis of V by inserting appropriate vectors into S.*

(d) *If W is a subspace of V, then $\dim(W) \le \dim(V)$. Moreover if $\dim(W) = \dim(V)$, then $W = V$.*

5 Dimensions of Subspaces

5.1 Row Space, Column Space and Nullspace

Definition 5.1 For an $m \times n$ matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ & & \cdots & \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix},$$

the vectors

$$\begin{aligned} \boldsymbol{r}_1 &= [a_{1,1}, a_{1,2}, \ldots, a_{1,n}] \\ \boldsymbol{r}_2 &= [a_{2,1}, a_{2,2}, \ldots, a_{2,n}] \\ \vdots & \quad \vdots \\ \boldsymbol{r}_m &= [a_{m,1}, a_{m,2}, \ldots, a_{m,n}] \end{aligned}$$

in $\boldsymbol{R}^n$ formed from the rows of A are called the *row vectors* of A and the vectors

$$\boldsymbol{c}_1 = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix}, \boldsymbol{c}_2 = \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix}, \cdots, \boldsymbol{c}_n = \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix}$$

in $\boldsymbol{R}^n$ formed from the columns of A are called the *column vectors* of A.

Definition 5.2 Let A be an $m \times n$ matrix, then the subspace of $\boldsymbol{R}^n$ spanned by the row vectors of A is called the *row space* of A, and the subspace of $\boldsymbol{R}^m$ spanned by the column vectors of A is called the *column* space of A. The solution space of the homogeneous system of equations $A\boldsymbol{x} = \boldsymbol{0}$, which is a subspace of $\boldsymbol{R}^n$, is called the *nullspace* of A.

The dimension of the column space of a matrix A is called the *rank* of A and is denoted by $\text{rank}(A)$. The dimension of the nullspace of A is called the *nullity* of A and is denoted by $\text{nullity}(A)$.

Proposition 5.1 (5.5.1) *A system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}$ is in the column space of A.*

Theorem 5.2 (5.5.2) *If $\boldsymbol{x}_0$ denotes any single solution of a consistent linear system $A\boldsymbol{x} = \boldsymbol{b}$, and if $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ form a basis for the nullspace of A, then every solution of $A\boldsymbol{x} = \boldsymbol{b}$ can be expressed in the form*

$$\boldsymbol{x} = \boldsymbol{x}_0 + c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k$$

and, conversely, for all choices of scalars $c_1, c_2, \cdots, c_k$, the vector $\boldsymbol{x}$ in this formula is a solution of $A\boldsymbol{x} = \boldsymbol{b}$.

Lemma 5.3 *Let A be an $m \times r$ matrix and B an $r \times n$ matrix. Then*

$$\mathcal{R}(AB) \subset \mathcal{R}(B),\ \mathcal{C}(AB) \subset \mathcal{C}(A).$$

Proposition 5.4 (5.5.3, 5.5.4, 5.5.5) *Let A be an $m \times n$ matrix and P an invertible matrix of size $m \times m$,*

(a) *Elementary row operations do not change the nullspace of a matrix. Moreover, $\mathcal{N}(A) = \mathcal{N}(PA)$.*

(b) *Elementary row operations do not change the row space of a matrix. Moreover, $\mathcal{R}(A) = \mathcal{R}(PA)$.*

(c) *$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\} \subset \boldsymbol{R}^n$ is a linearly independent set if and only if $\{P\boldsymbol{v}_1, P\boldsymbol{v}_2, \ldots, P\boldsymbol{v}_r\} \subset \boldsymbol{R}^n$ is a linearly independent set.*

5.2 Rank and Nullity

Proposition 5.5 (5.5.6) *If a matrix R is in row-echelon form, then the row vectors with the leading 1's form a basis for the row space of R, and the column vectors with the leading 1's of the row vectors form a basis for the column space of R.*

Theorem 5.6 (5.6.1) *If A is any matrix, then the row space and column space of A have the same dimension. Hence* $\text{rank}(A) = \text{rank}(A^T)$.

Theorem 5.7 (5.6.3) *If A is a matrix with n columns, then*

$$\text{rank}(A) + \text{nullity}(A) = n.$$

6 Inner Product Spaces

6.1 Inner Product, Norm and Distance

Definition 6.1 An *inner product* on a real vector space V is a function that associates a real number $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ with each pair of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V in such a way that the following axioms are satisfied for all vectors $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{z}$ in V and all scalars k.

(a) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$ (Symmetry axiom)

(b) $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{z} \rangle = \langle \boldsymbol{u}, \boldsymbol{z} \rangle + \langle \boldsymbol{v}, \boldsymbol{z} \rangle$ (Additive axiom)

(c) $\langle k\boldsymbol{u}, \boldsymbol{v} \rangle = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ (Homogeneity axiom)

(d) $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ (Positivity axiom)
and if $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}$.

A real vector space with an inner product is called a *real inner product space.*

Definition 6.2 An *inner product* on a complex vector space V is a function that associates a real number $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ with each pair of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V in such a way that the following axioms are satisfied for all vectors $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{z}$ in V and all scalars k ($k \in \boldsymbol{C}$).

(a) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}$ (Symmetry axiom)

(b) $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{z} \rangle = \langle \boldsymbol{u}, \boldsymbol{z} \rangle + \langle \boldsymbol{v}, \boldsymbol{z} \rangle$ (Additive axiom)

(c) $\langle k\boldsymbol{u}, \boldsymbol{v} \rangle = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ (Homogeneity axiom)

(d) $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ (Positively axiom)

and if $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ if and only if $\boldsymbol{v} = \mathbf{0}$.

A real vector space with an inner product is called a *real inner product space*.

Definition 6.3 If V is an inner product space, then the *norm* (or *length*) of a vector $\boldsymbol{u} \in V$ is denoted by $\|\boldsymbol{u}\|$ and is defined by

$$\|\boldsymbol{u}\| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle^{1/2}.$$

The *distance* between two points (vectors) $\boldsymbol{u}$ and $\boldsymbol{v}$ is denoted by $d(\boldsymbol{u}, \boldsymbol{v})$ and is defined by

$$d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|.$$

6.2 Properties of Inner Product Space

The following inequality is called the Cauchy-Schwarz Inequality.

Theorem 6.1 (6.2.1) *If $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in a real inner product space, then*

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|.$$

Equality holds if and only if $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly dependent.

Theorem 6.2 (6.2.2) *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in an inner product space V, and k a scalar. Then:*

(a) $\|\boldsymbol{u}\| \geq 0$.

(b) $\|\boldsymbol{u}\| = 0$ *if and only if* $\boldsymbol{u} = \mathbf{0}$.

(c) $\|k\boldsymbol{u}\| = |k|\|\boldsymbol{u}\|$.

(d) $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. *(Triangle inequality)*

Theorem 6.3 (6.2.3) *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in an inner product space V, and k a scalar. Then:*

(a) $d(\boldsymbol{u}, \boldsymbol{v}) \geq 0$.

(b) $d(\boldsymbol{u}, \boldsymbol{v}) = 0$ *if and only if* $\boldsymbol{u} = \boldsymbol{v}$.

(c) $d(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{v}, \boldsymbol{u})$.

(d) $d(\boldsymbol{u}, \boldsymbol{v}) \leq d(\boldsymbol{u}, \boldsymbol{w}) + d(\boldsymbol{w}, \boldsymbol{v})$. *(Triangle inequality)*

Theorem 6.4 (6.2.4) *If $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in an inner vector space, then*

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 \Leftrightarrow \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0.$$

7 Orthogonal Bases

7.1 Gram-Schmidt Proceess

Definition 7.1 A set of vectors in an inner product space is called an *orthogonal set* if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is called *orhonormal*.

Proposition 7.1 *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be an orthogonal set of nonzero vectors in an inner product space.*

(a) *S is a linearly independent set.*

(b) *Let $W = \mathrm{Span}(S)$ and $\boldsymbol{w} \in W$, then*

$$\boldsymbol{w} = \frac{\langle \boldsymbol{w}, \boldsymbol{v}_1 \rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{w}, \boldsymbol{v}_2 \rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{w}, \boldsymbol{v}_m \rangle}{\|\boldsymbol{v}_m\|^2}\boldsymbol{v}_m.$$

(c) *If $\boldsymbol{v} \in V$, then*

$$\langle \boldsymbol{v} - \mathrm{proj}_W(\boldsymbol{v}), \boldsymbol{w} \rangle = 0 \textit{ for all } \boldsymbol{w} \in W.$$

where $\mathrm{proj}_W(\boldsymbol{v})$ is defined by the following.

$$\frac{\langle \boldsymbol{v}, \boldsymbol{v}_1 \rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_2 \rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_m \rangle}{\|\boldsymbol{v}_m\|^2}\boldsymbol{v}_m.$$

Definition 7.2 Let W be a subspace of an inner product space V. A vector $\boldsymbol{u}$ in V is said to be *orthogonal to* W if it is orthogonal to every vector in W, and the set of all vectors in V that are orthogonal to W is called the *orthogonal complement* of W and is denoted by $W^\perp$. Hence $W^\perp = \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0 \text{ for all } \boldsymbol{w} \in W\}$.

Theorem 7.2 ((6.3.6) Gram-Schmidt Process) *Every nonzero finite-dimensional inner product space has an orthonormal basis.*

Theorem 7.3 (6.2.5, 6.2.6, 6.3.4) *If W is a subspace of a finite-dimensional inner product space V, then*

(a) *$W^\perp$ is a subspace of V.*

(b) *The only vector common to W and $W^\perp$ is $\mathbf{0}$.*

(c) $(W^\perp)^\perp = W$.

(d) $\dim V = \dim W + \dim W^\perp$.

(e) *Every vector $\boldsymbol{v} \in V$ is expressed as a sum $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u}$ such that $\boldsymbol{w} \in W$ and $\boldsymbol{u} \in W^\perp$.*

8 General Linear Transformations

8.1 Basic Properties

Definition 8.1 If $T : V \to W$ is a function from a vector space V into a vector space W, then T is called a *linear transformation* from V to W if , for all vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V and all scalars c,

(a) $T(\boldsymbol{u}+\boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ (b) $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$.

In the special case where $V = W$, the linear transformation $T : V \to V$ is called a *linear operator* of V.

Example 8.1 A linear transformation from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$ is first defined in Definition 2.2 as a function

$$T(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$$

for which the equations relating $y_1, y_2, \ldots, y_m$ with $x_1, x_2, \ldots, x_n$ are linear, and it was expressed by a matrix multiplication:

$$T(\boldsymbol{x}) = A\boldsymbol{x}, \text{ where } A = [T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), \ldots, T(\boldsymbol{e}_n)].$$

The matrix A was called the standard matrix and denoted by $A = [T]$ and $T = T_A$. Moreover, linear transformations were characterized by the two properties in Definition 8.1. See Theorem 2.2.

Example 8.2 [Examples 11, 12] Let $C^\infty(a, b)$ be the set of functions that are differentiable for all degrees of differentiation

1. $D : C^\infty(a, b) \to C^\infty(a, b)$ $(f(x) \mapsto f'(x))$ is a linear operator.
2. $I : C^\infty(a, b) \to C^\infty(a, b)$ $(f(x) \mapsto \int_a^x f(t)dt)$ is a linear operator.

Lemma 8.1 (8.1.1) *If $T : V \to W$ is a linear transformation, then*

(a) $T(\boldsymbol{0}) = \boldsymbol{0}$.

(b) $T(-\boldsymbol{v}) = -T(\boldsymbol{v})$ *for all* $\boldsymbol{v} \in V$.

(c) $T\left(\sum_{i=1}^m k_i \boldsymbol{v}_i\right) = \sum_{i=1}^m k_i T(\boldsymbol{v}_i)$ *for all* $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m \in V$ *and all scalars* $k_1, k_2, \ldots, k_m$.

Proposition 8.2 *Let T_1 and T_2 be linear transformations from V to W, and $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis of V*[1]*. Then the following are equivalent.*

(a) $T_1 = T_2$, *i.e.,* $T_1(\boldsymbol{v}) = T_2(\boldsymbol{v})$ *for all* $\boldsymbol{v} \in V$.

(b) $T_1(\boldsymbol{v}_i) = T_2(\boldsymbol{v}_i)$ *for all* $i = 1, 2, \ldots, n$.

[1]The condition $V = \mathrm{Span}(S)$ is enough.

Proposition 8.3 *Let V and W be vector spaces, $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis of V and $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n \in W$. Then there exists a unique linear transformation $T : V \to W$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for all $i = 1, 2, \ldots, n$.*

Proposition 8.4 (8.1.2) *Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations. Then the composition of T_2 with T_1 defined by*

$$T_2 \circ T_1 : U \to W \ (\boldsymbol{x} \mapsto T_2(T_1(\boldsymbol{x}))).$$

is a linear transformation.

8.2 Kernel and Range

Proposition 8.5 (8.2.1) *If $T : V \to W$ is a linear transformation, then*

(a) $\{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\}$ *is a subspace of* V.

(b) $\{T(\boldsymbol{v}) \mid \boldsymbol{v} \in V\}$ *is a subspace of* W.

Definition 8.2 If $T : V \to W$ is a linear transformation, then the set of vectors in V that T maps into $\boldsymbol{0}$ is called the *kernel* of T; it is denoted by $\mathrm{Ker}(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the *range* of T; it is denoted by $\mathrm{Im}(T)$. The dimension of the range of T is called the *rank* of T and is denoted by $\mathrm{rank}(T)$, the dimension of the kernel is called the *nullity* of T and is denoted by $\mathrm{nullity}(T)$.

$\mathrm{Ker}(T) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\} \subset V$, $\mathrm{Im}(T) = \{T(\boldsymbol{v}) \mid \boldsymbol{v} \in V\} \subset W$, and $\mathrm{nullity}(T) = \dim(\mathrm{Ker}(T))$, $\mathrm{rank}(T) = \dim(\mathrm{Im}(T))$.

The following is a generalization of Theorem 5.7. See also Theorem 7.3.

Theorem 8.6 (8.2.3) *If $T : V \to W$ is a linear transformation from an n-dimensional vector space V to a vector space W, then*

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = n.$$

Proposition 8.7 (8.3.1) *If $T : V \to W$ is a linear transformation, then the following are equivalent.*

(a) *T is one-to-one, i.e., injective.*

(b) $\mathrm{Ker}(T) = \{\boldsymbol{0}\}$.

(c) $\mathrm{nullity}(T) = 0$

Proposition 8.8 (8.3.2) *If V is a finite-dimensional vector space, and $T : V \to V$ is a linear operator, then the following are equivalent.*

(a) *T is one-to-one, i.e., injective.*

(b) $\mathrm{Ker}(T) = \{\boldsymbol{0}\}$.

(c) $\mathrm{nullity}(T) = 0$.

(d) *The range of T is V, i.e., surjective.*

(e) $\mathrm{rank}(T) = \dim V$.

9 Matrices and Linear Transformations

Definition 9.1 Suppose that V is an n-dimensional vector space with a basis $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and W is an m-dimensional vector space with a basis $B' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$. For $\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n \in V$, the vector $[x_1, x_2, \ldots, x_n]^T \in \boldsymbol{R}^n$ is called the *coordinate vector of* $\boldsymbol{x}$ *with respect to the basis* B and denoted by $[\boldsymbol{x}]_B$. Similarly for $\boldsymbol{y} = y_1\boldsymbol{w}_1 + y_2\boldsymbol{w}_2 + \cdots + y_m\boldsymbol{w}_m$, $[\boldsymbol{y}]_{B'} = [y_1, y_2, \ldots, y_m]^T \in \boldsymbol{R}^m$ is the coordinate vector of $\boldsymbol{y}$ with respect to the basis B'.

Let T be a linear transformation from V to W. Then the $m \times n$ matrix A defined by

$$A = [[T(\boldsymbol{v}_1)]_{B'}, [T(\boldsymbol{v}_2)]_{B'}, \ldots, [T(\boldsymbol{v}_m)]_{B'}]$$

is called the *matrix for* T *with respect to the bases* B *and* B' and denoted by $[T]_{B',B}$.

When $V = W$ and $B = B'$, we write $[T]_B$ for $[T]_{B,B}$ and $[T]_B$ is called the *matrix for* T *with respect to the basis* B.

Proposition 9.1 *Under the notation in Definition 9.1 the following hold.*

(a) $[T]_{B',B}[\boldsymbol{x}]_B = [T(\boldsymbol{x})]_{B'}$.

(b) $[T]_B[\boldsymbol{x}]_B = [T(\boldsymbol{x})]_B$, *with* $V = W$.

Proposition 9.2 (8.4.2) *Let* $T_1 : U \to V$ *and* $T_2 : V \to W$ *be linear transformations and* B, B' *and* B'' *basis of* U, V, *and* W *respectively. Then*

$$[T_2 \circ T_1]_{B'',B} = [T_2]_{B'',B'}[T_1]_{B',B}.$$

Proposition 9.3 (8.4.3) *Let* $T : V \to V$ *be a linear transformation. If* B *is a basis of* V, *then the following are equaivalent:*

(a) T *is one-to-one.*

(b) $[T]_B$ *is invertible.*

Morover, when these equivalent conditions hold,

$$[T^{-1}]_B = [T]_B^{-1}.$$

Theorem 9.4 (8.5.2) *Let* $T : V \to V$ *be a linear operator on a finite-dimensional vector space* V, *and let* B *and* B' *be bases for* V. *Then*

$$[T]_{B'} = P^{-1}[T]_B P$$

where P *is the transition matrix from* B' *to* B.

Definition 9.2 If A and B are square matrices, we say that B *is similar to* A if there is an invertible matrix P such that $B = P^{-1}AP$.

Example 9.1 Let $V = \boldsymbol{R}^3$. In Quizzes we showed that V has three bases. $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, $B' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{e}_1\}$, and $B'' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$, where

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}, \ \boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 7 \\ 4 \end{bmatrix}, \ \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ -8 \\ -6 \end{bmatrix},$$

$$\boldsymbol{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{-3}{\sqrt{14}} \\ \frac{-2}{\sqrt{14}} \end{bmatrix}, \ \boldsymbol{u}_2 = \begin{bmatrix} \frac{3}{\sqrt{70}} \\ \frac{5}{\sqrt{70}} \\ \frac{-6}{\sqrt{70}} \end{bmatrix}, \ \boldsymbol{u}_3 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{1}{\sqrt{5}} \end{bmatrix}.$$

The first is the standard basis, and the last is an orthonormal basis. Let $T = \text{proj}_U$.

By Quiz 7-5, since $T(\boldsymbol{e}_1)$, $T(\boldsymbol{e}_2)$, $T(\boldsymbol{e}_3)$ are

$$\frac{1}{5}[1, 0, -2]^T, \ [0, 1, 0]^T, \ \frac{1}{5}[-2, 0, 4]^T,$$

$$[T]_B = \begin{bmatrix} \frac{1}{5} & 0 & -\frac{2}{5} \\ 0 & 1 & 0 \\ -\frac{2}{5} & 0 & \frac{4}{5} \end{bmatrix}.$$

Similarly, $[T]'_B$ and $[T]_{B''}$ are

$$\begin{bmatrix} 1 & 0 & \frac{7}{5} \\ 0 & 1 & \frac{3}{5} \\ 0 & 0 & 0 \end{bmatrix}, \ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

As for $[I]_{B',B}$, $[I]_{B'',B}$, $[I]_{B'',B'}$, we have as follows:

$$\begin{bmatrix} 1 & -2 & 1 \\ -3 & 7 & 0 \\ -2 & 4 & 0 \end{bmatrix}, \ \begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{3}{\sqrt{70}} & \frac{2}{\sqrt{5}} \\ \frac{-3}{\sqrt{14}} & \frac{5}{\sqrt{70}} & 0 \\ \frac{-2}{\sqrt{14}} & \frac{-6}{\sqrt{70}} & \frac{1}{\sqrt{5}} \end{bmatrix},$$

$$\begin{bmatrix} \frac{\sqrt{14}}{14} & \frac{31\sqrt{70}}{70} & \frac{-7\sqrt{5}}{10} \\ 0 & \frac{\sqrt{70}}{5} & \frac{-3\sqrt{5}}{10} \\ 0 & 0 & \frac{\sqrt{5}}{2} \end{bmatrix}.$$

Recall the situation in Definition 9.1. Suppose

$$T(\boldsymbol{v}_i) = \sum_{j=1}^{m} a_{j,i}\boldsymbol{w}_j = a_{1,i}\boldsymbol{w}_1 + a_{2,i}\boldsymbol{w}_2 + \cdots + a_{m,i}\boldsymbol{w}_m.$$

Then $[T(\boldsymbol{v}_i)]_{B'} = [a_{1,i}, a_{2,i}, \ldots, a_{m,i}]^T$. Hence the ij entry of $[T]_{B,B'}$ is $a_{i,j}$.

Definition 9.3 An *isomorphism* between V and W is a bijective linear transformation form V to W. When there is an isomorphism between V and W, we say V and W are isomorphic.

When $V = W$, isomorphisms are called *automorphisms.*

Proposition 9.5 *Let* V *and* W *be finite-dimensional vector space. Then* V *and* W *are isomorphic, i.e., there is a bijective linear transformation from* V *to* W *if and only if* $\dim V = \dim W$. *In particular, every real vector space of dimension* n *is isomorphic to* $\boldsymbol{R}^n$.