Preface to the Second Edition

Let me begin by thanking the readers of the first edition for their many helpful comments and suggestions. The second edition represents a major change from the first edition. Indeed, one might say that it is a totally new book, with the exception of the general range of topics covered.

The text has been completely rewritten. I hope that an additional 12 years and roughly 20 books worth of experience has enabled me to improve the quality of my exposition. Also, the exercise sets have been completely rewritten.

The second edition contains two new chapters: a chapter on convexity, separation and positive solutions to linear systems (Chapter 15) and a chapter on the QR decomposition, singular values and pseudoinverses (Chapter 17). The treatments of tensor products and the umbral calculus have been greatly expanded and I have included discussions of determinants (in the chapter on tensor products), the complexification of a real vector space, Schur's lemma and Geršgorin disks.

Steven Roman *Irvine, California February 2005*

Preface to the First Edition

This book is a thorough introduction to linear algebra, for the graduate or advanced undergraduate student. Prerequisites are limited to a knowledge of the basic properties of matrices and determinants. However, since we cover the basics of vector spaces and linear transformations rather rapidly, a prior course in linear algebra (even at the sophomore level), along with a certain measure of "mathematical maturity," is highly desirable.

Chapter 0 contains a summary of certain topics in modern algebra that are required for the sequel. *This chapter should be skimmed quickly and then used primarily as a reference*. Chapters 1–3 contain a discussion of the basic properties of vector spaces and linear transformations.

Chapter 4 is devoted to a discussion of modules, emphasizing a comparison between the properties of modules and those of vector spaces. Chapter 5 provides more on modules. The main goals of this chapter are to prove that any two bases of a free module have the same cardinality and to introduce noetherian modules. However, the instructor may simply skim over this chapter, omitting all proofs. Chapter 6 is devoted to the theory of modules over a principal ideal domain, establishing the cyclic decomposition theorem for finitely generated modules. This theorem is the key to the structure theorems for finite-dimensional linear operators, discussed in Chapters 7 and 8.

Chapter 9 is devoted to real and complex inner product spaces. The emphasis here is on the finite-dimensional case, in order to arrive as quickly as possible at the finite-dimensional spectral theorem for normal operators, in Chapter 10. However, we have endeavored to state as many results as is convenient for vector spaces of arbitrary dimension.

The second part of the book consists of a collection of independent topics, with the one exception that Chapter 13 requires Chapter 12. Chapter 11 is on metric vector spaces, where we describe the structure of symplectic and orthogonal geometries over various base fields. Chapter 12 contains enough material on metric spaces to allow a unified treatment of topological issues for the basic

Hilbert space theory of Chapter 13. The rather lengthy proof that every metric space can be embedded in its completion may be omitted.

Chapter 14 contains a brief introduction to tensor products. In order to motivate the universal property of tensor products, without getting too involved in categorical terminology, we first treat both free vector spaces and the familiar direct sum, in a universal way. Chapter 15 [Chapter 16 in the second edition] is on affine geometry, emphasizing algebraic, rather than geometric, concepts.

The final chapter provides an introduction to a relatively new subject, called the umbral calculus. This is an algebraic theory used to study certain types of polynomial functions that play an important role in applied mathematics. We give only a brief introduction to the subject – emphasizing the algebraic aspects, rather than the applications. This is the first time that this subject has appeared in a true textbook.

One final comment. Unless otherwise mentioned, omission of a proof in the text is a tacit suggestion that the reader attempt to supply one.

Steven Roman *Irvine, California*

Chapter 2
Linear Transformations

Linear Transformations

Loosely speaking, a linear transformation is a function from one vector space to another that *preserves* the vector space operations. Let us be more precise.

Definition *Let* V *and* W *be vector spaces over a field* F. *A function* $\tau: V \to W$ *is a* **linear transformation** *if*

$$\tau(ru + sv) = r\tau(u) + s\tau(v)$$

for all scalars $r, s \in F$ *and vectors* $u, v \in V$. *A linear transformation* $\tau: V \to V$ *is called a* **linear operator** *on* V. *The set of all linear transformations from* V *to* W *is denoted by* $\mathcal{L}(V, W)$ *and the set of all linear operators on* V *is denoted by* $\mathcal{L}(V)$. ☐

We should mention that some authors use the term linear operator for any linear transformation from V to W.

Definition *The following terms are also employed:*

1) **homomorphism** *for linear transformation*
2) **endomorphism** *for linear operator*
3) **monomorphism** *(or* **embedding***) for injective linear transformation*
4) **epimorphism** *for surjective linear transformation*
5) **isomorphism** *for bijective linear transformation.*
6) **automorphism** *for bijective linear operator.* ☐

Example 2.1

1) The derivative $D: V \to V$ is a linear operator on the vector space V of all infinitely differentiable functions on $\mathbb{R}$.

2) The integral operator $\tau: F[x] \to F[x]$ defined by
$$\tau(f) = \int_0^x f(t)dt$$
is a linear operator on $F[x]$.
3) Let A be an $m \times n$ matrix over F. The function $\tau_A: F^n \to F^m$ defined by $\tau_A(v) = Av$, where all vectors are written as column vectors, is a linear transformation from F^n to F^m. This function is just multiplication by A.
4) The coordinate map $\phi: V \to F^n$ of an n-dimensional vector space is a linear transformation from V to F^n. □

The set $\mathcal{L}(V, W)$ is a vector space in its own right and $\mathcal{L}(V)$ has the structure of an algebra, as defined in Chapter 0.

Theorem 2.1
1) The set $\mathcal{L}(V, W)$ is a vector space under ordinary addition of functions and scalar multiplication of functions by elements of F.
2) If $\sigma \in \mathcal{L}(U, V)$ and $\tau \in \mathcal{L}(V, W)$ then the composition $\tau\sigma$ is in $\mathcal{L}(U, W)$.
3) If $\tau \in \mathcal{L}(V, W)$ is bijective then $\tau^{-1} \in \mathcal{L}(W, V)$.
4) The vector space $\mathcal{L}(V)$ is an algebra, where multiplication is composition of functions. The identity map $\iota \in \mathcal{L}(V)$ is the multiplicative identity and the zero map $0 \in \mathcal{L}(V)$ is the additive identity.

Proof. We prove only part 3). Let $\tau: V \to W$ be a bijective linear transformation. Then $\tau^{-1}: W \to V$ is a well-defined function and since any two vectors w_1 and w_2 in W have the form $w_1 = \tau(v_1)$ and $w_2 = \tau(v_2)$, we have
$$\begin{aligned}\tau^{-1}(aw_1 + bw_2) &= \tau^{-1}(a\tau(v_1) + b\tau(v_2)) \\ &= \tau^{-1}(\tau(av_1 + bv_2)) \\ &= av_1 + bv_2 \\ &= a\tau^{-1}(w_1) + b\tau^{-1}(w_2)\end{aligned}$$
which shows that τ^{-1} is linear. □

One of the easiest ways to define a linear transformation is to give its values on a basis. The following theorem says that we may assign these values arbitrarily and obtain a unique linear transformation by linear extension to the entire domain.

Theorem 2.2 *Let V and W be vector spaces and let $\mathcal{B} = \{v_i \mid i \in I\}$ be a basis for V. Then we can define a linear transformation $\tau \in \mathcal{L}(V, W)$ by specifying the values of $\tau(v_i) \in W$* arbitrarily *for all $v_i \in \mathcal{B}$ and extending the domain of τ to V using linearity, that is,*
$$\tau(a_1v_1 + \cdots + a_nv_n) = a_1\tau(v_1) + \cdots + a_n\tau(v_n)$$

This process uniquely *defines a linear transformation, that is, if* $\tau, \sigma \in \mathcal{L}(V, W)$ *satisfy* $\tau(v_i) = \sigma(v_i)$ *for all* $v_i \in \mathcal{B}$ *then* $\tau = \sigma$.
Proof. The crucial point is that the extension by linearity is well-defined, since each vector in V has a unique representation as a linear combination of a finite number of vectors in $\mathcal{B}$. We leave the details to the reader. □

Note that if $\tau \in \mathcal{L}(V, W)$ and if S is a subspace of V, then the restriction $\tau|_S$ of τ to S is a linear transformation from S to W.

The Kernel and Image of a Linear Transformation

There are two very important vector spaces associated with a linear transformation τ from V to W.

Definition *Let* $\tau \in \mathcal{L}(V, W)$. *The subspace*

$$\ker(\tau) = \{v \in V \mid \tau(v) = 0\}$$

is called the **kernel** *of* τ *and the subspace*

$$\operatorname{im}(\tau) = \{\tau(v) \mid v \in V\}$$

is called the **image** *of* τ. *The dimension of* $\ker(\tau)$ *is called the* **nullity** *of* τ *and is denoted by* $\operatorname{null}(\tau)$. *The dimension of* $\operatorname{im}(\tau)$ *is called the* **rank** *of* τ *and is denoted by* $\operatorname{rk}(\tau)$. □

It is routine to show that $\ker(\tau)$ is a subspace of V and $\operatorname{im}(\tau)$ is a subspace of W. Moreover, we have the following.

Theorem 2.3 *Let* $\tau \in \mathcal{L}(V, W)$. *Then*
1) τ *is surjective if and only if* $\operatorname{im}(\tau) = W$
2) τ *is injective if and only if* $\ker(\tau) = \{0\}$
Proof. The first statement is merely a restatement of the definition of surjectivity. To see the validity of the second statement, observe that

$$\tau(u) = \tau(v) \Leftrightarrow \tau(u - v) = 0 \Leftrightarrow u - v \in \ker(\tau)$$

Hence, if $\ker(\tau) = \{0\}$ then $\tau(u) = \tau(v) \Leftrightarrow u = v$, which shows that τ is injective. Conversely, if τ is injective and $u \in \ker(\tau)$ then $\tau(u) = \tau(0)$ and so $u = 0$. This shows that $\ker(\tau) = \{0\}$. □

Isomorphisms

Definition *A bijective linear transformation* $\tau: V \to W$ *is called an* **isomorphism** *from* V *to* W. *When an isomorphism from* V *to* W *exists, we say that* V *and* W *are* **isomorphic** *and write* $V \approx W$. □

Example 2.2 Let $\dim(V) = n$. For any ordered basis $\mathcal{B}$ of V, the coordinate map $\phi_{\mathcal{B}}: V \to F^n$ that sends each vector $v \in V$ to its coordinate matrix

$[v]_{\mathcal{B}} \in F^n$ is an isomorphism. Hence, any n-dimensional vector space over F is isomorphic to F^n. □

Isomorphic vector spaces share many properties, as the next theorem shows. If $\tau \in \mathcal{L}(V, W)$ and $S \subseteq V$ we write

$$\tau(S) = \{\tau(s) \mid s \in S\}$$

Theorem 2.4 *Let $\tau \in \mathcal{L}(V, W)$ be an isomorphism. Let $S \subseteq V$. Then*
1) S spans V if and only if $\tau(S)$ spans W.
2) S is linearly independent in V if and only if $\tau(S)$ is linearly independent in W.
3) S is a basis for V if and only if $\tau(S)$ is a basis for W. □

An isomorphism can be characterized as a linear transformation $\tau: V \to W$ that maps a basis for V to a basis for W.

Theorem 2.5 *A linear transformation $\tau \in \mathcal{L}(V, W)$ is an isomorphism if and only if there is a basis $\mathcal{B}$ of V for which $\tau(\mathcal{B})$ is a basis of W. In this case, τ maps any basis of V to a basis of W.* □

The following theorem says that, up to isomorphism, there is only one vector space of any given dimension.

Theorem 2.6 *Let V and W be vector spaces over F. Then $V \approx W$ if and only if* $\dim(V) = \dim(W)$. □

In Example 2.2, we saw that any n-dimensional vector space is isomorphic to F^n. Now suppose that B is a set of cardinality κ and let $(F^B)_0$ be the vector space of all functions from B to F with finite support. We leave it to the reader to show that the functions $\delta_b \in (F^B)_0$ defined for all $b \in B$, by

$$\delta_b(x) = \begin{cases} 1 & \text{if } x = b \\ 0 & \text{if } x \neq b \end{cases}$$

form a basis for $(F^B)_0$, called the **standard basis**. Hence, $\dim((F^B)_0) = |B|$.

It follows that for any cardinal number κ, there is a vector space of dimension κ. Also, any vector space of dimension κ is isomorphic to $(F^B)_0$.

Theorem 2.7 *If n is a natural number then any n-dimensional vector space over F is isomorphic to F^n. If κ is any cardinal number and if B is a set of cardinality κ then any κ-dimensional vector space over F is isomorphic to the vector space $(F^B)_0$ of all functions from B to F with finite support.* □

The Rank Plus Nullity Theorem

Let $\tau \in \mathcal{L}(V, W)$. Since any subspace of V has a complement, we can write

$$V = \ker(\tau) \oplus \ker(\tau)^c$$

where $\ker(\tau)^c$ is a complement of $\ker(\tau)$ in V. It follows that

$$\dim(V) = \dim(\ker(\tau)) + \dim(\ker(\tau)^c)$$

Now, the restriction of τ to $\ker(\tau)^c$

$$\tau^c: \ker(\tau)^c \to W$$

is injective, since

$$\ker(\tau^c) = \ker(\tau) \cap \ker(\tau)^c = \{0\}$$

Also, $\operatorname{im}(\tau^c) \subseteq \operatorname{im}(\tau)$. For the reverse inclusion, if $\tau(v) \in \operatorname{im}(\tau)$ then since $v = u + w$ for $u \in \ker(\tau)$ and $w \in \ker(\tau)^c$, we have

$$\tau(v) = \tau(u) + \tau(w) = \tau(w) = \tau^c(w) \in \operatorname{im}(\tau^c)$$

Thus $\operatorname{im}(\tau^c) = \operatorname{im}(\tau)$. It follows that

$$\ker(\tau)^c \approx \operatorname{im}(\tau)$$

From this, we deduce the following theorem.

Theorem 2.8 *Let* $\tau \in \mathcal{L}(V, W)$.

1) *Any complement of* $\ker(\tau)$ *is isomorphic to* $\operatorname{im}(\tau)$

2) **(The rank plus nullity theorem)**

$$\dim(\ker(\tau)) + \dim(\operatorname{im}(\tau)) = \dim(V)$$

or, in other notation,

$$\operatorname{rk}(\tau) + \operatorname{null}(\tau) = \dim(V)$$ □

Theorem 2.8 has an important corollary.

Corollary 2.9 *Let* $\tau \in \mathcal{L}(V, W)$*, where* $\dim(V) = \dim(W) < \infty$*. Then* τ *is injective if and only if it is surjective.* □

Note that this result fails if the vector spaces are not finite-dimensional.

Linear Transformations from F^n to F^m

Recall that for any $m \times n$ matrix A over F the multiplication map

$$\tau_A(v) = Av$$

is a linear transformation. In fact, any linear transformation $\tau \in \mathcal{L}(F^n, F^m)$ has this form, that is, τ is just multiplication by a matrix, for we have

$$\big(\tau(e_1) \mid \cdots \mid \tau(e_n)\big)e_i = \big(\tau(e_1) \mid \cdots \mid \tau(e_n)\big)^{(i)} = \tau(e_i)$$

and so $\tau = \tau_A$ where

$$A = \big(\tau(e_1) \mid \cdots \mid \tau(e_n)\big)$$

Theorem 2.10
1) If A is an $m \times n$ matrix over F then $\tau_A \in \mathcal{L}(F^n, F^m)$.
2) If $\tau \in \mathcal{L}(F^n, F^m)$ then $\tau = \tau_A$ where

$$A = (\tau(e_1) \mid \cdots \mid \tau(e_n))$$

The matrix A is called the **matrix** *of τ.* □

Example 2.3 Consider the linear transformation $\tau: F^3 \to F^3$ defined by

$$\tau(x, y, z) = (x - 2y, z, x + y + z)$$

Then we have, in column form

$$\tau \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-2y \\ z \\ x+y+z \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and so the standard matrix of τ is

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

If $A \in \mathcal{M}_{m,n}$ then since the image of τ_A is the column space of A, we have

$$\dim(\ker(\tau_A)) + \mathrm{rk}(A) = \dim(F^n)$$

This gives the following useful result.

Theorem 2.11 *Let A be an $m \times n$ matrix over F.*
1) $\tau_A: F^n \to F^m$ is injective if and only if $\mathrm{rk}(A) = n$.
2) $\tau_A: F^n \to F^m$ is surjective if and only if $\mathrm{rk}(A) = m$. □

Change of Basis Matrices

Suppose that $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C} = (c_1, \ldots, c_n)$ are ordered bases for a vector space V. It is natural to ask how the coordinate matrices $[v]_{\mathcal{B}}$ and $[v]_{\mathcal{C}}$ are related. The map that takes $[v]_{\mathcal{B}}$ to $[v]_{\mathcal{C}}$ is $\phi_{\mathcal{B},\mathcal{C}} = \phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}$ and is called the **change of basis operator** (or **change of coordinates operator**). Since $\phi_{\mathcal{B},\mathcal{C}}$ is an operator on F^n, it has the form τ_A where

$$\begin{aligned} A &= (\phi_{\mathcal{B},\mathcal{C}}(e_1), \ldots, \phi_{\mathcal{B},\mathcal{C}}(e_n)) \\ &= (\phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}([b_1]_{\mathcal{B}}), \ldots, \phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}([b_n]_{\mathcal{B}})) \\ &= ([b_1]_{\mathcal{C}}, \ldots, [b_n]_{\mathcal{C}})) \end{aligned}$$

We denote A by $M_{\mathcal{B},\mathcal{C}}$ and call it the **change of basis matrix** from $\mathcal{B}$ to $\mathcal{C}$.

Theorem 2.12 *Let $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C}$ be ordered bases for a vector space V. Then the change of basis operator $\phi_{\mathcal{B},\mathcal{C}} = \phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}$ is an automorphism of F^n, whose standard matrix is*

$$M_{\mathcal{B},\mathcal{C}} = ([b_1]_{\mathcal{C}}, \ldots, [b_n]_{\mathcal{C}}))$$

Hence

$$[v]_{\mathcal{C}} = M_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$$

and $M_{\mathcal{C},\mathcal{B}} = M_{\mathcal{B},\mathcal{C}}^{-1}$. □

Consider the equation

$$A = M_{\mathcal{B},\mathcal{C}}$$

or equivalently,

$$A = ([b_1]_{\mathcal{C}}, \ldots, [b_n]_{\mathcal{C}}))$$

Then given any two of A (an invertible $n \times n$ matrix), $\mathcal{B}$ (an ordered basis for F^n) and $\mathcal{C}$ (an order basis for F^n), the third component is uniquely determined by this equation. This is clear if $\mathcal{B}$ and $\mathcal{C}$ are given or if A and $\mathcal{C}$ are given. If A and $\mathcal{B}$ are given then there is a unique $\mathcal{C}$ for which $A^{-1} = M_{\mathcal{C},\mathcal{B}}$ and so there is a unique $\mathcal{C}$ for which $A = M_{\mathcal{B},\mathcal{C}}$.

Theorem 2.13 *If we are given any two of the following:*
1) An invertible $n \times n$ matrix A.
2) An ordered basis $\mathcal{B}$ for F^n.
3) An ordered basis $\mathcal{C}$ for F^n.
then the third is uniquely determined by the equation

$$A = M_{\mathcal{B},\mathcal{C}}$$ □

The Matrix of a Linear Transformation

Let $\tau: V \to W$ be a linear transformation, where $\dim(V) = n$ and $\dim(W) = m$ and let $\mathcal{B} = (b_1, \ldots, b_n)$ be an ordered basis for V and $\mathcal{C}$ an ordered basis for W. Then the map

$$\theta: [v]_{\mathcal{B}} \to [\tau(v)]_{\mathcal{C}}$$

is a *representation* of τ as a linear transformation from F^n to F^m, in the sense that knowing θ (along with $\mathcal{B}$ and $\mathcal{C}$, of course) is equivalent to knowing τ. Of course, this representation depends on the choice of ordered bases $\mathcal{B}$ and $\mathcal{C}$.

Since θ is a linear transformation from F^n to F^m, it is just multiplication by an $m \times n$ matrix A, that is

$$[\tau(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}}$$

Indeed, since $[b_i]_{\mathcal{B}} = e_i$, we get the columns of A as follows:

$$A^{(i)} = Ae_i = A[v]_{\mathcal{B}} = [\tau(b_i)]_{\mathcal{C}}$$

Theorem 2.14 *Let $\tau \in \mathcal{L}(V,W)$ and let $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C}$ be ordered bases for V and W, respectively. Then τ can be represented with respect to $\mathcal{B}$ and $\mathcal{C}$ as matrix multiplication, that is*

$$[\tau(v)]_{\mathcal{C}} = [\tau]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$$

where

$$[\tau]_{\mathcal{B},\mathcal{C}} = ([\tau(b_1)]_{\mathcal{C}} \mid \cdots \mid [\tau(b_n)]_{\mathcal{C}})$$

is called the **matrix of τ with respect to the bases** *$\mathcal{B}$ and $\mathcal{C}$. When $V = W$ and $\mathcal{B} = \mathcal{C}$, we denote $[\tau]_{\mathcal{B},\mathcal{B}}$ by $[\tau]_{\mathcal{B}}$ and so*

$$[\tau(v)]_{\mathcal{B}} = [\tau]_{\mathcal{B}}[v]_{\mathcal{B}}$$ □

Example 2.4 Let $D\colon \mathcal{P}_2 \to \mathcal{P}_2$ be the derivative operator, defined on the vector space of all polynomials of degree at most 2. Let $\mathcal{B} = \mathcal{C} = (1, x, x^2)$. Then

$$[D(1)]_{\mathcal{C}} = [0]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, [D(x)]_{\mathcal{C}} = [1]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, [D(x^2)]_{\mathcal{C}} = [2x]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

and so

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, for example, if $p(x) = 5 + x + 2x^2$ then

$$[Dp(x)]_{\mathcal{C}} = [D]_{\mathcal{B}}\,[p(x)]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}$$

and so $Dp(x) = 1 + 4x$. □

The following result shows that we may work equally well with linear transformations or with the matrices that represent them (with respect to fixed ordered bases $\mathcal{B}$ and $\mathcal{C}$). This applies not only to addition and scalar multiplication, but also to matrix multiplication.

Theorem 2.15 *Let V and W be vector spaces over F, with ordered bases $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C} = (c_1, \ldots, c_m)$, respectively.*

1) *The map $\mu: \mathcal{L}(V, W) \to \mathcal{M}_{m,n}(F)$ defined by*
$$\mu(\tau) = [\tau]_{\mathcal{B},\mathcal{C}}$$
is an isomorphism and so $\mathcal{L}(V, W) \approx \mathcal{M}_{m,n}(F)$.
2) *If $\sigma \in \mathcal{L}(U, V)$ and $\tau \in \mathcal{L}(V, W)$ and if $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ are ordered bases for U, V and W, respectively then*
$$[\tau\sigma]_{\mathcal{B},\mathcal{D}} = [\tau]_{\mathcal{C},\mathcal{D}}[\sigma]_{\mathcal{B},\mathcal{C}}$$
Thus, the matrix of the product (composition) $\tau\sigma$ is the product of the matrices of τ and σ. In fact, this is the primary motivation for the definition of matrix multiplication.

Proof. To see that μ is linear, observe that for all i
$$\begin{aligned}[s\sigma + t\tau]_{\mathcal{B},\mathcal{C}}[b_i]_{\mathcal{B}} &= [(s\sigma + t\tau)(b_i)]_{\mathcal{C}} \\ &= [s\sigma(b_i) + t\tau(b_i)]_{\mathcal{C}} \\ &= s[\sigma(b_i)]_{\mathcal{C}} + t[\tau(b_i)]_{\mathcal{C}} \\ &= s[\sigma]_{\mathcal{B},\mathcal{C}}[b_i]_{\mathcal{B}} + t[\tau]_{\mathcal{B},\mathcal{C}}[b_i]_{\mathcal{B}} \\ &= (s[\sigma]_{\mathcal{B},\mathcal{C}} + t[\tau]_{\mathcal{B},\mathcal{C}})[b_i]_{\mathcal{B}}\end{aligned}$$
and since $[b_i]_{\mathcal{B}} = e_i$ is a standard basis vector, we conclude that
$$[s\sigma + t\tau]_{\mathcal{B},\mathcal{C}} = s[\sigma]_{\mathcal{B},\mathcal{C}} + t[\tau]_{\mathcal{B},\mathcal{C}}$$
and so μ is linear. If $A \in \mathcal{M}_{m,n}$, we define τ by the condition $[\tau(b_i)]_{\mathcal{C}} = A^{(i)}$, whence $\mu(\tau) = A$ and μ is surjective. Since $\dim(\mathcal{L}(V, W)) = \dim(\mathcal{M}_{m,n}(F))$, the map μ is an isomorphism. To prove part 2), we have
$$[\tau\sigma]_{\mathcal{B},\mathcal{D}}[v]_{\mathcal{B}} = [\tau(\sigma(v))]_{\mathcal{D}} = [\tau]_{\mathcal{C},\mathcal{D}}[\sigma(v)]_{\mathcal{C}} = [\tau]_{\mathcal{C},\mathcal{D}}[\sigma]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} \qquad \square$$

Change of Bases for Linear Transformations

Since the matrix $[\tau]_{\mathcal{B},\mathcal{C}}$ that represents τ depends on the ordered bases $\mathcal{B}$ and $\mathcal{C}$, it is natural to wonder how to choose these bases in order to make this matrix as simple as possible. For instance, can we always choose the bases so that τ is represented by a diagonal matrix?

As we will see in Chapter 7, the answer to this question is no. In that chapter, we will take up the general question of how best to represent a linear operator by a matrix. For now, let us take the first step and describe the relationship between the matrices $[\tau]_{\mathcal{B},\mathcal{C}}$ and $[\tau]_{\mathcal{B}',\mathcal{C}'}$ of τ with respect to two different pairs $(\mathcal{B}, \mathcal{C})$ and $(\mathcal{B}', \mathcal{C}')$ of ordered bases. Multiplication by $[\tau]_{\mathcal{B}',\mathcal{C}'}$ sends $[v]_{\mathcal{B}'}$ to

$[\tau(v)]_{\mathcal{C}'}$. This can be reproduced by first switching from $\mathcal{B}'$ to $\mathcal{B}$, then applying $[\tau]_{\mathcal{B},\mathcal{C}}$ and finally switching from $\mathcal{C}$ to $\mathcal{C}'$, that is,

$$[\tau]_{\mathcal{B}',\mathcal{C}'} = M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B}',\mathcal{B}} = M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B},\mathcal{B}'}^{-1}$$

Theorem 2.16 *Let $\tau \in \mathcal{L}(V,W)$ and let $(\mathcal{B},\mathcal{C})$ and $(\mathcal{B}',\mathcal{C}')$ be pairs of ordered bases of V and W, respectively. Then*

$$[\tau]_{\mathcal{B}',\mathcal{C}'} = M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B}',\mathcal{B}} \tag{2.1}$$ □

When $\tau \in \mathcal{L}(V)$ is a linear operator on V, it is generally more convenient to represent τ by matrices of the form $[\tau]_{\mathcal{B}}$, where the ordered bases used to represent vectors in the domain and image are the same. When $\mathcal{B} = \mathcal{C}$, Theorem 2.16 takes the following important form.

Corollary 2.17 *Let $\tau \in \mathcal{L}(V)$ and let $\mathcal{B}$ and $\mathcal{C}$ be ordered bases for V. Then the matrix of τ with respect to $\mathcal{C}$ can be expressed in terms of the matrix of τ with respect to $\mathcal{B}$ as follows*

$$[\tau]_{\mathcal{C}} = M_{\mathcal{B},\mathcal{C}}[\tau]_{\mathcal{B}}M_{\mathcal{B},\mathcal{C}}^{-1} \tag{2.2}$$ □

Equivalence of Matrices

Since the change of basis matrices are precisely the invertible matrices, (2.1) has the form

$$[\tau]_{\mathcal{B}',\mathcal{C}'} = P[\tau]_{\mathcal{B},\mathcal{C}}Q^{-1}$$

where P and Q are invertible matrices. This motivates the following definition.

Definition *Two matrices A and B are* **equivalent** *if there exist invertible matrices P and Q for which*

$$B = PAQ^{-1}$$ □

We remarked in Chapter 0 that B is equivalent to A if and only if B can be obtained from A by a series of elementary row and column operations. Performing the row operations is equivalent to multiplying the matrix A on the left by P and performing the column operations is equivalent to multiplying A on the right by Q^{-1}.

In terms of (2.1), we see that performing row operations (premultiplying by P) is equivalent to changing the basis used to represent vectors in the image and performing column operations (postmultiplying by Q^{-1}) is equivalent to changing the basis used to represent vectors in the domain.

According to Theorem 2.16, if A and B are matrices that represent τ with respect to possibly different ordered bases then A and B are equivalent. The converse of this also holds.

Theorem 2.18 *Let V and W be vector spaces with $\dim(V) = n$ and $\dim(W) = m$. Then two $m \times n$ matrices A and B are equivalent if and only if they represent the same linear transformation $\tau \in \mathcal{L}(V,W)$, but possibly with respect to different ordered bases. In this case, A and B represent exactly the same set of linear transformations in $\mathcal{L}(V,W)$.*
Proof. If A and B represent τ, that is, if

$$A = [\tau]_{\mathcal{B},\mathcal{C}} \text{ and } B = [\tau]_{\mathcal{B}',\mathcal{C}'}$$

for ordered bases $\mathcal{B}, \mathcal{C}, \mathcal{B}'$ and $\mathcal{C}'$ then Theorem 2.16 shows that A and B are equivalent. Now suppose that A and B are equivalent, say

$$B = PAQ^{-1}$$

where P and Q are invertible. Suppose also that A represents a linear transformation $\tau \in \mathcal{L}(V,W)$ for some ordered bases $\mathcal{B}$ and $\mathcal{C}$, that is,

$$A = [\tau]_{\mathcal{B},\mathcal{C}}$$

Theorem 2.13 implies that there is a unique ordered basis $\mathcal{B}'$ for V for which $Q = M_{\mathcal{B},\mathcal{B}'}$ and a unique ordered basis $\mathcal{C}'$ for W for which $P = M_{\mathcal{C},\mathcal{C}'}$. Hence

$$B = M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B}',\mathcal{B}} = [\tau]_{\mathcal{B}',\mathcal{C}'}$$

Hence, B also represents τ. By symmetry, we see that A and B represent the same set of linear transformations. This completes the proof. □

We remarked in Example 0.3 that every matrix is equivalent to exactly one matrix of the block form

$$J_k = \begin{bmatrix} I_k & 0_{k,n-k} \\ 0_{m-k,k} & 0_{m-k,n-k} \end{bmatrix}_{\text{block}}$$

Hence, the set of these matrices is a set of canonical forms for equivalence. Moreover, the rank is a complete invariant for equivalence. In other words, two matrices are equivalent if and only if they have the same rank.

Similarity of Matrices

When a linear operator $\tau \in \mathcal{L}(V)$ is represented by a matrix of the form $[\tau]_{\mathcal{B}}$, equation (2.2) has the form

$$[\tau]_{\mathcal{B}'} = P[\tau]_{\mathcal{B}}P^{-1}$$

where P is an invertible matrix. This motivates the following definition.

Definition *Two matrices A and B are* **similar** *if there exists an invertible matrix P for which*

$$B = PAP^{-1}$$

The equivalence classes associated with similarity are called **similarity classes**. □

The analog of Theorem 2.18 for square matrices is the following.

Theorem 2.19 *Let V be a vector space of dimension n. Then two $n \times n$ matrices A and B are similar if and only if they represent the same linear operator $\tau \in \mathcal{L}(V)$, but possibly with respect to different ordered bases. In this case, A and B represent exactly the same set of linear operators in $\mathcal{L}(V)$.*

Proof. If A and B represent $\tau \in \mathcal{L}(V)$, that is, if

$$A = [\tau]_{\mathcal{B}} \text{ and } B = [\tau]_{\mathcal{C}}$$

for ordered bases $\mathcal{B}$ and $\mathcal{C}$ then Corollary 2.17 shows that A and B are similar. Now suppose that A and B are similar, say

$$B = PAP^{-1}$$

Suppose also that A represents a linear operator $\tau \in \mathcal{L}(V)$ for some ordered basis $\mathcal{B}$, that is,

$$A = [\tau]_{\mathcal{B}}$$

Theorem 2.13 implies that there is a unique ordered basis $\mathcal{C}$ for V for which $P = M_{\mathcal{B},\mathcal{C}}$. Hence

$$B = M_{\mathcal{B},\mathcal{C}}[\tau]_{\mathcal{B}}M_{\mathcal{B},\mathcal{C}}^{-1} = [\tau]_{\mathcal{C}}$$

Hence, B also represents τ. By symmetry, we see that A and B represent the same set of linear operators. This completes the proof. □

We will devote much effort in Chapter 7 to finding a canonical form for similarity.

Similarity of Operators

We can also define similarity of operators.

Definition *Two linear operators $\tau, \sigma \in \mathcal{L}(V)$ are* **similar** *if there exists an automorphism $\phi \in \mathcal{L}(V)$ for which*

$$\sigma = \phi\tau\phi^{-1}$$

The equivalence classes associated with similarity are called **similarity classes**. □

The analog of Theorem 2.19 in this case is the following.

Theorem 2.20 *Let V be a vector space of dimension n. Then two linear operators τ and σ on V are similar if and only if there is a matrix $A \in \mathcal{M}_n$ that represents both operators (but with respect to possibly different ordered bases). In this case, τ and σ are represented by exactly the same set of matrices in $\mathcal{M}_n$.*
Proof. If τ and σ are represented by $A \in \mathcal{M}_n$, that is, if

$$[\tau]_{\mathcal{B}} = A = [\sigma]_{\mathcal{C}}$$

for ordered bases $\mathcal{B}$ and $\mathcal{C}$ then

$$[\sigma]_{\mathcal{C}} = [\tau]_{\mathcal{B}} = M_{\mathcal{C},\mathcal{B}}[\tau]_{\mathcal{C}}M_{\mathcal{B},\mathcal{C}}$$

Let $\phi \in \mathcal{L}(V)$ be the automorphism of V defined by $\phi(c_i) = b_i$, where $\mathcal{B} = \{b_1, \ldots, b_n\}$ and $\mathcal{C} = \{c_1, \ldots, c_n\}$. Then

$$[\phi]_{\mathcal{C}} = ([\phi(c_1)]_{\mathcal{C}} \mid \cdots \mid [\phi(c_n)]_{\mathcal{C}}) = ([b_1]_{\mathcal{C}} \mid \cdots \mid [b_n]_{\mathcal{C}}) = M_{\mathcal{B},\mathcal{C}}$$

and so

$$[\sigma]_{\mathcal{C}} = [\phi]_{\mathcal{C}}^{-1}[\tau]_{\mathcal{C}}[\phi]_{\mathcal{C}} = [\phi^{-1}\tau\phi]_{\mathcal{C}}$$

from which it follows that σ and τ are similar. Conversely, suppose that τ and σ are similar, say

$$\sigma = \phi\tau\phi^{-1}$$

Suppose also that τ is represented by the matrix $A \in \mathcal{M}_n$, that is,

$$A = [\tau]_{\mathcal{B}}$$

for some ordered basis $\mathcal{B}$. Then

$$[\sigma]_{\mathcal{B}} = [\phi\tau\phi^{-1}]_{\mathcal{B}} = [\phi]_{\mathcal{B}}[\tau]_{\mathcal{B}}[\phi]_{\mathcal{B}}^{-1}$$

If we set $c_i = \phi(b_i)$ then $\mathcal{C} = (c_1, \ldots, c_n)$ is an ordered basis for V and

$$[\phi]_{\mathcal{B}} = ([\phi(b_1)]_{\mathcal{B}} \mid \cdots \mid [\phi(b_n)]_{\mathcal{B}}) = ([c_1]_{\mathcal{B}} \mid \cdots \mid [c_n]_{\mathcal{B}}) = M_{\mathcal{C},\mathcal{B}}$$

Hence

$$[\sigma]_{\mathcal{B}} = M_{\mathcal{C},\mathcal{B}}[\tau]_{\mathcal{B}}M_{\mathcal{C},\mathcal{B}}^{-1}$$

It follows that

$$A = [\tau]_{\mathcal{B}} = M_{\mathcal{B},\mathcal{C}}[\sigma]_{\mathcal{B}}M_{\mathcal{B},\mathcal{C}}^{-1} = [\sigma]_{\mathcal{C}}$$

and so A also represents σ. By symmetry, we see that τ and σ are represented by the same set of matrices. This completes the proof. $\square$

Invariant Subspaces and Reducing Pairs

The restriction of a linear operator $\tau \in \mathcal{L}(V)$ to a subspace S of V is not necessarily a linear operator on S. This prompts the following definition.

Definition *Let $\tau \in \mathcal{L}(V)$. A subspace S of V is said to be* **invariant under** τ *or* τ-**invariant** *if $\tau(S) \subseteq S$, that is, if $\tau(s) \in S$ for all $s \in S$. Put another way, S is invariant under τ if the restriction $\tau|_S$ is a linear operator on S.* $\square$

If

$$V = S \oplus T$$

then the fact that S is τ-invariant does not imply that the complement T is also τ-invariant. (The reader may wish to supply a simple example with $V = \mathbb{R}^2$.)

Definition *Let $\tau \in \mathcal{L}(V)$. If $V = S \oplus T$ and if both S and T are τ-invariant, we say that the pair (S,T)* **reduces** τ. $\square$

A reducing pair can be used to decompose a linear operator into a direct sum as follows.

Definition *Let $\tau \in \mathcal{L}(V)$. If (S,T) reduces τ we write*

$$\tau = \tau|_S \oplus \tau|_T$$

and call τ the **direct sum** *of $\tau|_S$ and $\tau|_T$. Thus, the expression*

$$\rho = \sigma \oplus \tau$$

means that there exist subspaces S and T of V for which (S,T) reduces ρ and

$$\sigma = \rho|_S \text{ and } \tau = \rho|_T \qquad \square$$

The concept of the direct sum of linear operators will play a key role in the study of the structure of a linear operator.

Topological Vector Spaces

This section is for readers with some familiarity with point-set topology. The **standard topology** on $\mathbb{R}^n$ is the topology for which the set of **open rectangles**

$$B = \{I_1 \times \cdots \times I_n \mid I_i\text{'s are open intervals in } \mathbb{R}\}$$

is a basis (in the sense of topology), that is, a subset of $\mathbb{R}^n$ is open if and only if it is a union of sets in B. The standard topology is the topology induced by the Euclidean metric on $\mathbb{R}^n$.

The standard topology on $\mathbb{R}^n$ has the property that the addition function

$$\mathcal{A}: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n: (v, w) \to v + w$$

and the scalar multiplication function

$$\mathcal{M}: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n: (r, v) \to rv$$

are continuous. As such, $\mathbb{R}^n$ is a *topological vector space*. Also, any linear functional $f: \mathbb{R}^n \to \mathbb{R}$ is a continuous map.

More generally, any real vector space V endowed with a topology $\mathcal{T}$ is called a **topological vector space** if the operations of addition $\mathcal{A}: V \times V \to V$ and scalar multiplication $\mathcal{M}: \mathbb{R} \times V \to V$ are continuous under $\mathcal{T}$.

Let V be a real vector space of dimension n and fix an ordered basis $\mathcal{B} = (v_1, \ldots, v_n)$ for V. Consider the coordinate map

$$\phi = \phi_{\mathcal{B}}: V \to \mathbb{R}^n: v \to [v]_{\mathcal{B}}$$

and its inverse

$$\psi_{\mathcal{B}} = \phi_{\mathcal{B}}^{-1}: \mathbb{R}^n \to V: (a_1, \ldots, a_n) \to \sum a_i v_i$$

We claim that there is precisely one topology $\mathcal{T} = \mathcal{T}_V$ on V for which V becomes a topological vector space and for which all linear functionals are continuous. This is called the **natural topology** on V. In fact, the natural topology is the topology for which $\phi_{\mathcal{B}}$ (and therefore also $\psi_{\mathcal{B}}$) is a homeomorphism, for any basis $\mathcal{B}$. (Recall that a **homeomorphism** is a bijective map that is continuous and has a continuous inverse.)

Once this has been established, it will follow that the open sets in $\mathcal{T}$ are precisely the images of the open sets in $\mathbb{R}^n$ under the map $\psi_{\mathcal{B}}$. A basis for the natural topology is given by

$$\begin{aligned}\{\psi_{\mathcal{B}}(I_1 \times \cdots \times I_n) \mid I_i\text{'s are open intervals in } \mathbb{R}\} \\ = \Big\{\sum_{r_i \in I_i} r_i v_i \mid I_i\text{'s are open intervals in } \mathbb{R}\Big\}\end{aligned}$$

First, we show that if V is a topological vector space under a topology $\mathcal{T}$ then ψ is continuous. Since $\psi = \sum \psi_i$ where $\psi_i: \mathbb{R}^n \to V$ is defined by

$$\psi_i(a_1, \ldots, a_n) = a_i v_i$$

it is sufficient to show that these maps are continuous. (The sum of continuous maps is continuous.) Let O be an open set in $\mathcal{T}$. Then

$$\mathcal{M}^{-1}(O) = \{(r, x) \in \mathbb{R} \times V \mid rx \in O\}$$

is open in $\mathbb{R} \times V$. We need to show that the set

$$\psi_i^{-1}(O) = \{(a_1, \ldots, a_n) \in \mathbb{R}^n \mid a_i v_i \in O\}$$

is open in $\mathbb{R}^n$, so let $(a_1, \ldots, a_n) \in \psi_i^{-1}(O)$. Thus, $a_i v_i \in O$. It follows that $(a_i, v_i) \in \mathcal{M}^{-1}(O)$, which is open, and so there is an open interval $I \subseteq \mathbb{R}$ and an open set $B \in \mathcal{T}$ of V for which

$$(a_i, v_i) \in I \times B \subseteq \mathcal{M}^{-1}(O)$$

Then the open set $U = \mathbb{R} \times \cdots \times \mathbb{R} \times I \times \mathbb{R} \times \cdots \times \mathbb{R}$, where the factor I is in the ith position, has the property that $\psi_i(U) \subseteq O$. Thus

$$(a_1, \ldots, a_n) \in U \subseteq \psi_i^{-1}(O)$$

and so $\psi_i^{-1}(O)$ is open. Hence, ψ_i, and therefore also ψ, is continuous.

Next we show that if every linear functional on V is continuous under a topology $\mathcal{T}$ on V then the coordinate map ϕ is continuous. If $v \in V$ denote by $[v]_{\mathcal{B},i}$ the ith coordinate of $[v]_{\mathcal{B}}$. The map $\mu\colon V \to \mathbb{R}$ defined by $\mu(v) = [v]_{\mathcal{B},i}$ is a linear functional and so is continuous by assumption. Hence, for any open interval $I_i \in \mathbb{R}$ the set

$$A_i = \{v \in V \mid [v]_{\mathcal{B},i} \in I_i\}$$

is open. Now, if I_i are open intervals in $\mathbb{R}$ then

$$\phi^{-1}(I_1 \times \cdots \times I_n) = \{v \in V \mid [v]_{\mathcal{B}} \in I_1 \times \cdots \times I_n\} = \bigcap A_i$$

is open. Thus, ϕ is continuous.

Thus, if a topology $\mathcal{T}$ has the property that V is a topological vector space and every linear functional is continuous, then ϕ and $\psi = \phi^{-1}$ are homeomorphisms. This means that $\mathcal{T}$, if it exists, must be unique.

It remains to prove that the topology $\mathcal{T}$ on V that makes ϕ a homeomorphism has the property that V is a topological vector space under $\mathcal{T}$ and that any linear functional f on V continuous.

As to addition, the maps $\phi\colon V \to \mathbb{R}^n$ and $(\phi \times \phi)\colon V \times V \to \mathbb{R}^n \times \mathbb{R}^n$ are homeomorphisms and the map $\mathcal{A}'\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and so the map $\mathcal{A}\colon V \times V \to V$, being equal to $\phi^{-1} \circ \mathcal{A}' \circ (\phi \times \phi)$, is also continuous.

As to scalar multiplication, the maps $\phi\colon V \to \mathbb{R}^n$ and $(\iota \times \phi)\colon \mathbb{R} \times V \to \mathbb{R} \times \mathbb{R}^n$ are homeomorphisms and the map $\mathcal{M}'\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and so the map $\mathcal{M}\colon V \times V \to V$, being equal to $\phi^{-1} \circ \mathcal{M}' \circ (\iota \times \phi)$, is also continuous.

Now let f be a linear functional. Since ϕ is continuous if and only if $f \circ \phi^{-1}$ is continuous, we can confine attention to $V = \mathbb{R}^n$. In this case, if $e_1, \ldots, e_n$ is the standard basis for $\mathbb{R}^n$ and $|f(e_i)| \le M$, then for any $x = (a_1, \ldots, a_n) \in \mathbb{R}^n$ we have

$$|f(x)| = \left|\sum a_i f(e_i)\right| \le \sum |a_i||f(e_i)| \le M \sum |a_i|$$

Now, if $|x| < \epsilon/Mn$ then $|a_i| < \epsilon/Mn$ and so $|f(x)| < \epsilon$, which implies that f is continuous.

According to the Riesz representation theorem and the Cauchy–Schwarz inequality, we have

$$\|f(x)\| \le \|\mathcal{R}_f\|\|x\|$$

Hence, $x_n \to 0$ implies $f(x_n) \to 0$ and so by linearity, $x_n \to x$ implies $f(x_n) \to x$ and so f is continuous.

Theorem 2.21 *Let V be a real vector space of dimension n. There is a unique topology on V, called the* **natural topology** *for which V is a topological vector space and for which all linear functionals on V are continuous. This topology is determined by the fact that the coordinate map $\phi: V \to \mathbb{R}^n$ is a homeomorphism.* □

Linear Operators on $V^{\mathbb{C}}$

A linear operator τ on a real vector space V can be extended to a linear operator $\tau^{\mathbb{C}}$ on the complexification $V^{\mathbb{C}}$ by defining

$$\tau^{\mathbb{C}}(u + vi) = \tau(u) + \tau(v)i$$

Here are the basic properties of this **complexification** of τ.

Theorem 2.22 *If $\tau, \sigma \in \mathcal{L}(V)$ then*
1) $(a\tau)^{\mathbb{C}} = a\tau^{\mathbb{C}}$, $a \in \mathbb{R}$
2) $(\tau + \sigma)^{\mathbb{C}} = \tau^{\mathbb{C}} + \sigma^{\mathbb{C}}$
3) $(\tau\sigma)^{\mathbb{C}} = \tau^{\mathbb{C}}\sigma^{\mathbb{C}}$
4) $[\tau(v)]^{\mathbb{C}} = \tau^{\mathbb{C}}(v^{\mathbb{C}})$. □

Let us recall that for any ordered basis $\mathcal{B}$ for V and any vector $v \in V$ we have

$$[v + 0i]_{\text{cpx}(\mathcal{B})} = [v]_{\mathcal{B}}$$

Now, if $\mathcal{B}$ is a basis for V, then the ith column of $[\tau]_{\mathcal{B}}$ is

$$[\tau(b_i)]_{\mathcal{B}} = [\tau(b_i) + 0i]_{\text{cpx}(\mathcal{B})} = [\tau^{\mathbb{C}}(b_i + 0i)]_{\text{cpx}(\mathcal{B})}$$

which is the ith column of the coordinate matrix of $\tau^{\mathbb{C}}$ with respect to the basis cpx($\mathcal{B}$). Thus we have the following theorem.

Theorem 2.23 *Let $\tau \in \mathcal{L}(V)$ where V is a real vector space. The matrix of $\tau^{\mathbb{C}}$ with respect to the basis* $\text{cpx}(\mathcal{B})$ *is equal to the matrix of τ with respect to the basis $\mathcal{B}$*

$$[\tau^{\mathbb{C}}]_{\text{cpx}(\mathcal{B})} = [\tau]_{\mathcal{B}}$$

Hence, if a real matrix A represents a linear operator τ on V then A also represents the complexification $\tau^{\mathbb{C}}$ of τ on $V^{\mathbb{C}}$. □

Exercises

1. Let $A \in \mathcal{M}_{m,n}$ have rank k. Prove that there are matrices $X \in \mathcal{M}_{m,k}$ and $Y \in \mathcal{M}_{k,n}$, both of rank k, for which $A = XY$. Prove that A has rank 1 if and only if it has the form $A = x^t y$ where x and y are row matrices.
2. Prove Corollary 2.9 and find an example to show that the corollary does not hold without the finiteness condition.
3. Let $\tau \in \mathcal{L}(V, W)$. Prove that τ is an isomorphism if and only if it carries a basis for V to a basis for W.
4. If $\tau \in \mathcal{L}(V_1, W_1)$ and $\sigma \in \mathcal{L}(V_2, W_2)$ we define the external direct sum $\tau \boxplus \sigma \in \mathcal{L}(V_1 \boxplus V_2, W_1 \boxplus W_2)$ by

$$(\tau \boxplus \sigma)((v_1, v_2)) = (\tau(v_1), \sigma(v_2))$$

   Show that $\tau \boxplus \sigma$ is a linear transformation.
5. Let $V = S \oplus T$. Prove that $S \oplus T \approx S \boxplus T$. Thus, internal and external direct sums are equivalent up to isomorphism.
6. Let $V = A + B$ and consider the external direct sum $E = A \boxplus B$. Define a map $\tau: A \boxplus B \to V$ by $\tau(v, w) = v + w$. Show that τ is linear. What is the kernel of τ? When is τ an isomorphism?
7. Let $\mathcal{T}$ be a subset of $\mathcal{L}(V)$. A subspace S of V is **$\mathcal{T}$-invariant** if S is τ-invariant for every $\tau \in \mathcal{T}$. Also, V is **$\mathcal{T}$-irreducible** if the only $\mathcal{T}$-invariant subspaces of V are $\{0\}$ and V. Prove the following form of *Schur's lemma.* Suppose that $\mathcal{T}_V \subseteq \mathcal{L}(V)$ and $\mathcal{T}_W \subseteq \mathcal{L}(W)$ and V is $\mathcal{T}_V$-irreducible and W is $\mathcal{T}_W$-irreducible. Let $\alpha \in \mathcal{L}(V, W)$ satisfy $\alpha\mathcal{T}_V = \mathcal{T}_W\alpha$, that is, for any $\mu \in \mathcal{T}_V$ there is a $\lambda \in \mathcal{T}_W$ such that $\alpha\mu = \lambda\alpha$. Prove that $\alpha = 0$ or α is an isomorphism.
8. Let $\tau \in \mathcal{L}(V)$ where $\dim(V) < \infty$. If $\text{rk}(\tau^2) = \text{rk}(\tau)$ show that $\text{im}(\tau) \cap \ker(\tau) = \{0\}$.
9. Let $\tau \in \mathcal{L}(U, V)$ and $\sigma \in \mathcal{L}(V, W)$. Show that

$$\text{rk}(\tau\sigma) \le \min\{\text{rk}(\tau), \text{rk}(\sigma)\}$$

10. Let $\tau \in \mathcal{L}(U, V)$ and $\sigma \in \mathcal{L}(V, W)$. Show that

$$\text{null}(\tau\sigma) \le \text{null}(\tau) + \text{null}(\sigma)$$

11. Let $\tau, \sigma \in \mathcal{L}(V)$ where τ is invertible. Show that

$$\text{rk}(\tau\sigma) = \text{rk}(\sigma\tau) = \text{rk}(\sigma)$$

12. Let $\tau, \sigma \in \mathcal{L}(V,W)$. Show that
$$\text{rk}(\tau+\sigma) \le \text{rk}(\tau) + \text{rk}(\sigma)$$
13. Let S be a subspace of V. Show that there is a $\tau \in \mathcal{L}(V)$ for which $\ker(\tau) = S$. Show also that there exists a $\sigma \in \mathcal{L}(V)$ for which $\text{im}(\sigma) = S$.
14. Suppose that $\tau, \sigma \in \mathcal{L}(V)$.
    a) Show that $\sigma = \tau\mu$ for some $\mu \in \mathcal{L}(V)$ if and only if $\text{im}(\sigma) \subseteq \text{im}(\tau)$.
    b) Show that $\sigma = \mu\tau$ for some $\mu \in \mathcal{L}(V)$ if and only if $\ker(\tau) \subseteq \ker(\sigma)$.
15. Let $V = S_1 \oplus S_2$. Define linear operators ρ_i on V by $\rho_i(s_1 + s_2) = s_i$ for $i = 1,2$. These are referred to as **projection operators**. Show that
    1) $\rho_i^2 = \rho_i$
    2) $\rho_1 + \rho_2 = I$, where I is the identity map on V.
    3) $\rho_i\rho_j = 0$ for $i \ne j$ where 0 is the zero map.
    4) $V = \text{im}(\rho_1) \oplus \text{im}(\rho_2)$
16. Let $\dim(V) < \infty$ and suppose that $\tau \in \mathcal{L}(V)$ satisfies $\tau^2 = 0$. Show that $2\text{rk}(\tau) \le \dim(V)$.
17. Let A be an $m \times n$ matrix over F. What is the relationship between the linear transformation $\tau_A\colon F^n \to F^m$ and the system of equations $AX = B$? Use your knowledge of linear transformations to state and prove various results concerning the system $AX = B$, especially when $B = 0$.
18. Let V have basis $\mathcal{B} = \{v_1, \ldots, v_n\}$. Suppose that for each $1 \le i, j \le n$ we define $\tau_{i,j} \in \mathcal{L}(V)$ by
$$\tau_{i,j}(v_k) = \begin{cases} v_k & \text{if } k \ne i \\ v_i + v_j & \text{if } k = i \end{cases}$$
    Prove that the $\tau_{i,j}$ are invertible and form a basis for $\mathcal{L}(V)$.
19. Let $\tau \in \mathcal{L}(V)$. If S is a τ-invariant subspace of V must there be a subspace T of V for which (S,T) reduces τ?
20. Find an example of a vector space V and a proper subspace S of V for which $V \approx S$.
21. Let $\dim(V) < \infty$. If τ, $\sigma \in \mathcal{L}(V)$ prove that $\sigma\tau = \iota$ implies that τ and σ are invertible and that $\sigma = p(\tau)$ for some polynomial $p(x) \in F[x]$.
22. Let $\tau \in \mathcal{L}(V)$ where $\dim(V) < \infty$. If $\tau\sigma = \sigma\tau$ for all $\sigma \in \mathcal{L}(V)$ show that $\tau = a\iota$, for some $a \in F$, where ι is the identity map.
23. Let $A, B \in \mathcal{M}_n(F)$. Let K be a field containing F. Show that if A and B are similar over K, that is, if $B = PAP^{-1}$ where $P \in \mathcal{M}_n(K)$ then A and B are also similar over F, that is, there exists $Q \in \mathcal{M}_n(F)$ for which $B = QAQ^{-1}$. *Hint*: consider the equation $XA - BX = 0$ as a homogeneous system of linear equations with coefficients in F. Does it have a solution? Where?
24. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a continuous function with the property that
$$f(x+y) = f(x) + f(y)$$
    Prove that f is a linear functional on $\mathbb{R}^n$.

25. Prove that any linear functional $f: \mathbb{R}^n \to \mathbb{R}$ is a continuous map.
26. Prove that any subspace S of $\mathbb{R}^n$ is a closed set or, equivalently, that $S^c = \mathbb{R}^n \setminus S$ is open, that is, for any $x \in S^c$ there is an open ball $B(s, \epsilon)$ centered at x with radius $\epsilon > 0$ for which $B(x, \epsilon) \subseteq S^c$.
27. Prove that any linear transformation $\tau: V \to W$ is continuous under the natural topologies of V and W.
28. Prove that any surjective linear transformation τ from V to W (both finite-dimensional topological vector spaces under the natural topology) is an open map, that is, τ maps open sets to open sets.
29. Prove that any subspace S of a finite-dimensional vector space V is a closed set or, equivalently, that S^c is open, that is, for any $x \in S^c$ there is an open ball $B(s, \epsilon)$ centered at x with radius $\epsilon > 0$ for which $B(x, \epsilon) \subseteq S^c$.
30. Let S be a subspace of V with $\dim(V) < \infty$.
    a) Show that the subspace topology on S inherited from V is the natural topology.
    b) Show that the natural topology on V/S is the topology for which the natural projection map $\pi: V \to V/S$ continuous and open.
31. If V is a real vector space then $V^{\mathbb{C}}$ is a complex vector space. Thinking of $V^{\mathbb{C}}$ as a vector space $(V^{\mathbb{C}})_{\mathbb{R}}$ over $\mathbb{R}$, show that $(V^{\mathbb{C}})_{\mathbb{R}}$ is isomorphic to the external direct product $V \boxplus V$.
34. (When is a complex linear map a complexification?) Let V be a real vector space with complexification $V^{\mathbb{C}}$ and let $\sigma \in \mathcal{L}(V^{\mathbb{C}})$. Prove that σ is a complexification, that is, σ has the form $\tau^{\mathbb{C}}$ for some $\tau \in \mathcal{L}(V)$ if and only if σ commutes with the conjugate map $\chi: V^{\mathbb{C}} \to V^{\mathbb{C}}$ defined by $\chi(u + iv) = u - iv$.
35. Let W be a complex vector space.
    a) Consider replacing the scalar multiplication on W by the operation

    $$(z, w) \to \overline{z}w$$

    where $z \in \mathbb{C}$ and $w \in W$. Show that the resulting set with the addition defined for the vector space W and with this scalar multiplication is a complex vector space, which we denote by $\overline{W}$.
    b) Show, without using dimension arguments, that $(W_{\mathbb{R}})^{\mathbb{C}} \approx W \boxplus \overline{W}$.
36. a) Let τ be a linear operator on the real vector space U with the property that $\tau^2 = -\iota$. Define a scalar multiplication on U by complex numbers as follows

    $$(a + bi) \cdot v = av + b\tau(v)$$

    for $a, b \in \mathbb{R}$ and $v \in U$. Prove that under the addition of U and this scalar multiplication U is a complex vector space, which we denote by U_τ.
    b) What is the relationship between U_τ and $V^{\mathbb{C}}$? Hint: consider $U = V \boxplus V$ and $\tau(u, v) = (-v, u)$.