Supplementary Notes on Linear Algebra

Mariusz Wodzicki

May 3, 2015
1 Vector spaces

1.1 Coordinatization of a vector space

1.1.1
Given a basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ in a vector space $V$, any vector $v \in V$ can be represented as a linear combination

$$v = \beta_1 b_1 + \cdots + \beta_n b_n$$

(1)

and this representation is unique, i.e., there is only one sequence of coefficients $\beta_1, \ldots, \beta_n$ for which (1) holds.

1.1.2
The correspondence between vectors in $V$ and the coefficients in the expansion (1) defines $n$ real valued functions on $V$,

$$b_i^*: V \to \mathbb{R}, \quad b_i^*(v) := \beta_i, \quad (i = 1, \ldots, n).$$

(2)

1.1.3
If

$$v' = \beta'_1 b_1 + \cdots + \beta'_n b_n$$

(3)

is another vector, then

$$v + v' = (\beta_1 + \beta'_1)b_1 + \cdots + (\beta_n + \beta'_n)b_n$$

which shows that

$$b_i^*(v + v') = \beta_i + \beta'_i = b_i^*(v) + b_i^*(v'), \quad (i = 1, \ldots, n).$$

Similarly, for any number $\alpha$, one has

$$\alpha v = (\alpha \beta_1) b_1 + \cdots + (\alpha \beta_n) b_n$$

which shows that

$$b_i^*(\alpha v) = \alpha b_i^*(v), \quad (i = 1, \ldots, n).$$
In other words, each function \( b^*_i : V \to \mathbb{R} \) is a linear transformation from the vector space \( V \) to the one-dimensional vector space \( \mathbb{R} \), and the correspondence

\[
v \mapsto [v]_B := \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}
\]

is a linear transformation from \( V \) to the \( n \)-dimensional vector space of column vectors \( \mathbb{R}^n \).

1.1.4 The coordinatization isomorphism of \( V \) with \( \mathbb{R}^n \)

The kernel of (4) is \{0\} since vectors \( b_1, \ldots, b_n \) are linearly independent. The range of (4) is the whole \( \mathbb{R}^n \) since vectors \( b_1, \ldots, b_n \) span \( V \). Thus, the correspondence \( v \mapsto [v]_B \) identifies \( V \) with the vector space \( \mathbb{R}^n \). We shall refer to (4) as the coordinatization of the vector space \( V \) in basis \( B \).

1.2 The dual space \( V^\vee \)

1.2.1

Linear transformations \( V \to \mathbb{R} \) are referred to as (linear) functionals on \( V \) (they are also called linear forms on \( V \)). Linear functionals form a vector space of their own which is called the dual of \( V \). We will denote it \( V^\vee \) (pronounce it “\( V \) dual” or “\( V \) check”).

1.2.2 An example: the trace of a matrix

The trace of an \( n \times n \) matrix \( A = [a_{ij}] \) is the sum of the diagonal entries,

\[
\text{tr} A := a_{11} + \cdots + a_{nn}.
\]

(5)

The correspondence \( A \mapsto \text{tr} A \) is a linear functional on the vector space Mat\(_n(\mathbb{R})\) of \( n \times n \) matrices.

Exercise 1 Calculate both \( \text{tr} AB \) and \( \text{tr} BA \) and show that

\[
\text{tr} AB = \text{tr} BA
\]

(6)

where

\[
A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{nn} \end{bmatrix}
\]

(7)
denotes an arbitrary \( m \times n \) matrix and
\[
B = \begin{bmatrix}
\beta_{11} & \cdots & \beta_{1m} \\
\vdots & \ddots & \vdots \\
\beta_{n1} & \cdots & \beta_{nm}
\end{bmatrix}
\quad (8)
\]
denotes an arbitrary \( n \times m \) matrix.

1.2.3 An example: the dual of the space of \( m \times n \) matrices

For any \( n \times m \) matrix (8), let us consider the linear functional on the space of \( m \times n \) matrices:
\[
\phi_B : A \mapsto \text{tr} \ AB \quad (A \in \text{Mat}_{mn}(\mathbb{R})).
\quad (9)
\]

Exercise 2 Calculate \( \phi_B(B^\tau) \), where \( B^\tau \) denotes the transpose of \( B \), and show that it vanishes if and only if \( B = 0 \). Deduce that \( \phi_B = 0 \) if and only if \( B = 0 \).

1.2.4

The correspondence
\[
\phi : \text{Mat}_{nm}(\mathbb{R}) \longrightarrow \text{Mat}_{mn}(\mathbb{R})^\vee, \quad B \mapsto \phi_B,
\quad (10)
\]
is a natural linear transformation from the space of \( n \times m \) matrices into the dual of the space of \( m \times n \) matrices. In view of Exercise 2 it is injective.

By considering bases in \( V \), in the next sections we will show that the dimension of \( V^\vee \) equals the dimension of \( V \) if the latter is finite. In particular, this will imply that \( \dim \text{Mat}_{nm}(\mathbb{R})^\vee = \dim \text{Mat}_{mn}(\mathbb{R}) \). Since the transposition of matrices,
\[
A \mapsto A^\tau \quad (A \in \text{Mat}_{mn}(\mathbb{R})),
\]
is an isomorphism of vector spaces, it will follow that
\[
\dim \text{Mat}_{nm}(\mathbb{R}) = \dim \text{Mat}_{mn}(\mathbb{R})^\vee.
\]
A corollary of this is that the dual space \( \text{Mat}_{mn}(\mathbb{R}) \) is naturally identified with the vector space of \( n \times m \) matrices, via identification (10).
1.2.5 The duality between the spaces of row and column vectors

In particular, the space of column vectors $\mathbb{R}^n = \text{Mat}_{n \times 1} (\mathbb{R})$ is naturally identified with the dual of the space of row vectors $\text{Mat}_{1 \times n} (\mathbb{R})$ and, vice-versa, the space of row vectors $\text{Mat}_{1 \times n} (\mathbb{R})$ is naturally identified with the dual of the space of column vectors $\mathbb{R}^n = \text{Mat}_{n \times 1} (\mathbb{R})$.

1.2.6 The coordinate functionals

The coordinatization isomorphism of $V$ with $\mathbb{R}^n$ is made up of $n$ coordinate functionals, cf. (2). They span $V^\ast$. Indeed, given a linear functional $\phi : V \to \mathbb{R}$, let

$$\alpha_1 := \phi(b_1), \ldots, \alpha_n := \phi(b_n).$$

Then, for any vector $v \in V$, one has

$$\phi(v) = \phi(\beta_1 b_1 + \cdots + \beta_n b_n)$$
$$= \beta_1 \phi(b_1) + \cdots + \beta_n \phi(b_n)$$
$$= \beta_1 \alpha_1 + \cdots + \beta_n \alpha_n$$
$$= \alpha_1 \beta_1 + \cdots + \alpha_n \beta_n$$
$$= \alpha_1 b_1^\ast(v) + \cdots + \alpha_n b_n^\ast(v)$$
$$= (\alpha_1 b_1^\ast + \cdots + \alpha_n b_n^\ast)(v)$$

which shows that the linear functional $\phi$ is a linear combination of functionals $b_1^\ast, \ldots, b_n^\ast$, $\phi = \alpha_1 b_1^\ast + \cdots + \alpha_n b_n^\ast$.

1.2.7

The coordinate functionals $b_1^\ast, \ldots, b_n^\ast$ are linearly independent. Indeed, if a linear combination

$$\alpha_1 b_1^\ast + \cdots + \alpha_n b_n^\ast$$

is the zero functional, then its values on $v = b_1, \ldots, b_n$ are all zero. But those values are:

$$\alpha_1, \ldots, \alpha_n,$$

since

$$b_i^\ast(b_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

(11)
1.2.8 The dual basis $B^*$

Thus, $B^* := \{b_1^*, \ldots, b_n^*\}$ forms a basis of the dual space. Note, that

$$\dim V^\vee = \dim V.$$ 

1.3 Scalar products

1.3.1 Bilinear pairings

A function of two vector arguments

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

is said to be a **bilinear pairing** on $V$ if it is a linear functional in each argument. (Bilinear pairings are also called **bilinear forms** on $V$.)

1.3.2

We say that the bilinear pairing is **nondegenerate** if, for any nonzero vector $v \in V$, there exists $v' \in V$, such that

$$\langle v, v' \rangle \neq 0.$$ 

1.3.3

We say that the bilinear pairing is **symmetric** if, for any vectors $v, v' \in V$, one has

$$\langle v', v \rangle = \langle v, v' \rangle.$$ 

1.3.4 Orthogonality

We say that vectors $v$ and $v'$ are **orthogonal** if $\langle v, v' \rangle = 0$. We denote this fact by $v \perp v'$.

1.3.5

If $X$ is a subset of $V$, the set of vectors orthogonal to every element of $X$ is denoted

$$X^\perp := \{v \in V \mid v \perp x \text{ for all } x \in X\}$$

**Exercise 3** Show that $X^\perp$ is a vector subspace of $V$ and

$$X \subseteq X^{\perp \perp}.$$
We say that the bilinear pairing is **positively defined** if, for any vector \( v \in V \), one has
\[
\langle v, v \rangle \geq 0.
\]

**Theorem 1.1 (The Cauchy-Schwarz Inequality)** Let \( \langle , \rangle \) be a positively defined symmetric bilinear pairing on a vector space \( V \). Then, for any vectors \( v, v' \in V \), one has the following inequality
\[
\langle v, v' \rangle^2 \leq \langle v, v \rangle \langle v', v' \rangle.
\]
(15)

**1.3.7**

We shall demonstrate (15) by considering the second degree polynomial
\[
p(t) \:= \langle tv + v', tv + v' \rangle = \langle v, v \rangle t^2 + (\langle v, v' \rangle + \langle v', v \rangle) t + \langle v', v' \rangle
\]
where
\[
a = \langle v, v \rangle, \quad b = 2\langle v, v' \rangle \quad \text{and} \quad c = \langle v', v' \rangle.
\]
In view of the hypothesis, \( p(t) \geq 0 \) for all real number \( t \). This is equivalent to the inequality \( b^2 \leq 4ac \) which yields inequality (15).

**1.3.8**

An immediate corollary of the Cauchy-Schwarz Inequality is that a symmetric bilinear pairing is nondegenerate and positively defined if and only if
\[
\langle v, v \rangle > 0
\]
for any nonzero vector in \( V \).

**1.3.9 Scalar products**

A nondegenerate positively defined symmetric bilinear pairing on \( V \) is called a **scalar product**.

**Exercise 4** Show that a set of nonzero vectors \( \{ v_1, \ldots, v_n \} \), mutually orthogonal with respect to some scalar product on \( V \), is linearly independent. (Hint: for a linear combination representing the zero vector,
\[
\alpha_1 v_1 + \cdots + \alpha_n v_n = 0
\]
calculate the scalar product of both sides with each \( v_i \).)
1.3.10 The associated norm

For any scalar product, the functional

\[ \langle v \rangle := \sqrt{\langle v, v \rangle} \quad (16) \]

is called the associated norm. Using the norm notation, we can rewrite the Cauchy-Schwarz Inequality as

\[ |\langle v, v' \rangle| \leq \|v\| \|v'\| \quad (17) \]

1.3.11 The Triangle Inequality

Note that

\[ \|v + v'\|^2 = \|v\|^2 + 2\langle v, v' \rangle + \|v'\|^2 \]

while

\[ (\|v\| + \|v'\|)^2 = \|v\|^2 + 2\|v\|\|v'\| + \|v'\|^2 \]

In view of the Cauchy-Schwarz Inequality, the bottom expression is not less than the top expression. Equivalently,

\[ \|v + v'\| \leq \|v\| + \|v'\| \quad (18) \]

for any pair of vectors \(v\) and \(v'\) in \(V\). This is known as the Triangle Inequality.

1.3.12

The associated norm satisfies also the following two conditions

\[ \|\alpha v\| = |\alpha| \|v\| \quad (19) \]

for any real number \(\alpha\) and any vector \(v \in V\), and

\[ \|v\| > 0 \quad (20) \]

for any nonzero vector \(v \in V\).

1.3.13 Norms on a vector space

Any function

\[ V \rightarrow [0, \infty] \]

that satisfies the Triangle Inequality (18) and conditions (19) and (20) is called a norm on \(V\).
1.3.14 Polarization Formula

In terms of the associated norm, the scalar product is expressed by means of the identity

$$\langle v, v' \rangle = \frac{1}{2} \left( \|v + v'\|^2 - \|v\|^2 - \|v'\|^2 \right).$$

(21)

known as the Polarization Formula. If a norm \( \| \| \) on a vector space \( V \) is associated with a scalar product, then the right-hand-side of (21) must depend on \( v \) linearly. If it does not, then that norm is not associated with a scalar product.

1.3.15 Quadratic forms

A function \( q: V \to \mathbb{R} \) is called a quadratic form if the pairing assigning the number

$$\langle v, v' \rangle := \frac{1}{2} \left( q(v + v') - q(v) - q(v') \right)$$

(22)

to a pair of vectors \( v \) and \( v' \) in \( V \), is bilinear. Note that the pairing defined by (22) is symmetric. Vice-versa, for any symmetric bilinear pairing \( \langle , \rangle \), the function

$$q(v) := \langle v, v \rangle \quad (v \in V),$$

(23)

is a quadratic form on \( V \).

1.3.16

We obtain a natural one-to-one correspondence between symmetric bilinear pairings and quadratic forms on \( V \)

$$\left\{ \text{symmetric bilinear pairings} \right\} \leftrightarrow \left\{ \text{quadratic forms} \right\}.$$

(24)

1.3.17

Nondegenerate symmetric pairings correspond to nondegenerate quadratic forms, i.e., the ones that satisfy

$$q(v) = 0 \quad \text{if and only if} \quad v = 0.$$
1.3.18
Positively defined symmetric bilinear pairings correspond to positively defined quadratic forms, i.e., the ones that satisfy

\[ q(v) \geq 0 \quad (v \in V). \]

1.3.19 **An example: the dot product**

Given a basis \( B = \{b_1, \ldots, b_n\} \) in \( V \), the dot product \( v \cdot_B v' \) of two vectors (1) and (3) is defined as

\[ v \cdot_B v' := \beta_1 \beta'_1 + \cdots + \beta_n \beta'_n. \tag{25} \]

It is the only scalar product on \( V \) for which \( B \) is orthonormal, i.e.,

\[ \langle b_i, b_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \tag{26} \]

1.3.20
In the special case of \( V = \mathbb{R}^n \) and \( B \) being the standard basis

\[ e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}, \tag{27} \]

we obtain the dot product on \( \mathbb{R}^n \).

1.3.21 **An example: the \( l_p \)-norms on \( \mathbb{R}^n \)**

For a positive number \( p > 0 \), consider the following functional on \( \mathbb{R}^n \),

\[ x \mapsto \|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}. \tag{28} \]

This functional satisfies the Triangle Inequality if and only if \( p \geq 1 \). For any \( p > 0 \), it satisfies the other two properties of a norm. Only for \( p = 2 \) the right-hand-side of the Polarization Formula is linear in \( v \). In that case, the scalar product is the dot product on \( \mathbb{R}^n \) and the \( l_2 \)-norm is known as the Euclidean norm. The vector space \( \mathbb{R}^n \) equipped with the \( l_2 \)-norm is referred to as the \( n \)-dimensional Euclidean space.
1.3.22 An example: the Killing scalar product

Consider the bilinear pairing on the vector space \( \text{Mat}_{mn}(\mathbb{R}) \) of \( m \times n \) matrices
\[
\langle A, B \rangle := \text{tr} A^\top B.
\] (29)

The pairing is known under the name of Killing form and plays a very important role in Representation Theory.

Exercise 5 Calculate \( \text{tr} A^\top A \) and show that \( \text{tr} A^\top A > 0 \) for all nonzero \( m \times n \) matrices. Explain why \( \text{tr} A^\top B = \text{tr} B^\top A \).

1.3.23

Note that the Killing scalar product on \( \text{Mat}_{n\times1}(\mathbb{R}) = \mathbb{R}^n \) coincides with the standard dot product on \( \mathbb{R}^n \).

1.3.24 Isometries

A linear transformation \( T: V \to V' \) between vector spaces equipped with bilinear pairings \( \langle , \rangle \) and, respectively, \( \langle , \rangle' \), is called an isometry if it preserves the value of the pairing, i.e., if
\[
\langle Tv_1, Tv_2 \rangle' = \langle v_1, v_2 \rangle
\]
for any pair of vectors \( v_1 \) and \( v_2 \) in \( V \).

1.3.25

The coordinatization isomorphism \( [ ]_B: V \to \mathbb{R}^n \) is an isometry between \( V \) equipped with the dot product \( \cdot_B \) and \( \mathbb{R}^n \) equipped with standard dot product:
\[
[v_1]_B \cdot [v_2]_B = v_1 \cdot_B v_2.
\] (30)

1.3.26 Description of bilinear pairings on a vector space with a basis

Given a basis \( \mathcal{B} = \{b_1, \ldots, b_n\} \) and an arbitrary bilinear pairing (12), let us consider the \( n \times n \) matrix \( Q = [q_{ij}] \) where
\[
q_{ij} := \langle b_i, b_j \rangle \quad (1 \leq i, j \leq n).
\] (31)
For a pair of vectors, as in (1) and (3), one has
\[
\langle v, v' \rangle = \langle \beta_1 b_1 + \cdots + \beta_n b_n, \beta'_1 b_1 + \cdots + \beta'_n b_n \rangle,
\]
\[
= \sum_{1 \leq i,j \leq n} \beta_i \langle b_i, b_j \rangle \beta'_j
\]
\[
= \sum_{1 \leq i,j \leq n} \beta_i q_{ij} \beta'_j \tag{32}
\]
\[
= [v]_B \cdot \left( Q [v']_B \right) \tag{33}
\]
\[
= (Q^T [v]_B) \cdot [v']_B \tag{34}
\]
where \( Q^T \) denotes the transpose matrix.

If we denote by \( \cdot_Q \) the bilinear pairing on \( \mathbb{R}^n \) given by
\[
x \cdot_Q y := x^T Q y, \tag{35}
\]
then (32) can be rewritten as
\[
\langle v, v' \rangle = [v]_B \cdot_Q [v']_B. \tag{36}
\]

1.3.27

We obtained a description of all bilinear pairings on a vector space with a chosen basis. They are in one-to-one correspondence with \( n \times n \) matrices. In particular, on \( \mathbb{R}^n \) every bilinear pairing is of the form (35) for a unique \( n \times n \) matrix \( Q \).

1.3.28

Nondegenerate pairings correspond to invertible matrices.

1.3.29

Symmetric pairings correspond to symmetric matrices.

1.3.30

Pairings for which the chosen basis is orthogonal correspond to diagonal matrices.
1.3.31
Positively defined pairings correspond to matrices $Q$ such that
\[ x^T Q x \geq 0 \]
for any $x \in \mathbb{R}^n$.

1.3.32 **Symplectic forms**

*Antisymmetric* pairings, i.e., the ones satisfying
\[ \langle v', v \rangle = -\langle v, v' \rangle \quad (v, v' \in V), \]
correspond to *antisymmetric matrices* $Q$, i.e., matrices whose transpose $Q^T$ equals $-Q$. Nondegenerate antisymmetric bilinear pairings are called **symplectic forms**. Vector spaces equipped with symplectic forms play a fundamental role in Physics, especially Mechanics, and in modern Mathematics.

1.3.33 **An example: the standard symplectic form on $\mathbb{R}^{2l}$**

For vectors
\[
\begin{align*}
x &= \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_l \\
\beta_1 \\
\vdots \\
\beta_l
\end{bmatrix} \\
x' &= \begin{bmatrix}
\alpha'_1 \\
\vdots \\
\alpha'_l \\
\beta'_1 \\
\vdots \\
\beta'_l
\end{bmatrix}
\end{align*}
\]
the formula
\[
\langle x, x' \rangle_{\text{sympl}} := (\alpha_1 \beta'_1 + \cdots + \alpha_l \beta'_l) - (\beta_1 \alpha'_1 + \cdots + \beta_l \alpha'_l)
\]
defines the *standard* symplectic form on $\mathbb{R}^{2l}$. This bilinear pairing corresponds to the $2l \times 2l$ matrix
\[
Q = \begin{bmatrix}
1 & \ddots & \ddots & 1 \\
-1 & 1 & \ddots & \ddots \\
& \ddots & \ddots & 1 \\
& & -1 & 1
\end{bmatrix}
\]
1.3.34
Note that the symplectic form (37) on $\mathbb{R}^2$ coincides with the determinant of a $2 \times 2$ matrix, viewed as a bilinear function of its columns.

1.3.35
One can equip any even dimensional vector space $V$ with a symplectic form, e.g., by coordinatizing $V$ and using the standard symplectic form (37). Symplectic forms do not exist on vector spaces of odd dimension.

1.3.36 The coordinatization isomorphism is an isometry
We observe that the coordinatization isomorphism is an isometry between $V$ equipped with a pairing $\langle, \rangle$ and $\mathbb{R}^n$ equipped with the pairing $\cdot \cdot$ defined in (35).

1.4 Linear transformations $V \rightarrow V^\vee$

1.4.1
Given a bilinear pairing $\langle, \rangle: V \times V \rightarrow \mathbb{R}$, the correspondence

$$v \mapsto v^* \in V^\vee,$$

where $v^*$ is the linear functional

$$v' \mapsto \langle v, v' \rangle \quad (v' \in V),$$

defines a linear transformation

$$^*: V \rightarrow V^\vee.$$  

(38)  

1.4.2
Vice-versa, any linear transformation $\Phi: V \rightarrow V^\vee$, produces a bilinear pairing

$$\langle v, v' \rangle := (\Phi(v))(v').$$

(40)
1.4.3

The two correspondences are mutually inverse. We thus obtain a natural one-to-one correspondence between bilinear pairings on $V$ and linear transformations $V \to V^\vee$,

$$\left\{ \text{bilinear pairings} \right\}_{\text{on } V} \longleftrightarrow \left\{ \text{linear transformations} \right\}_{V \to V^\vee}. \quad (41)$$

1.4.4

If $\langle \cdot , \cdot \rangle = \cdot_B$ is the dot product associated with a basis $\mathcal{B}$ in $V$, then $b_i^*$ in the sense of (38) coincide with the coordinate functionals introduced in Section 1.1.2. Note that the linear transformation $V \to V^\vee$ that corresponds to the scalar product $\cdot_B$, sends basis vectors $b_1, \ldots, b_n$ in $V$ to the basis vectors $b_1^*, \ldots, b_n^*$ in $V^\vee$.

1.4.5

The pairing is nondegenerate if and only if the kernel of the induced transformation $V \to V^\vee$ is zero. For finite-dimensional spaces the latter occurs precisely when $V \to V^\vee$ is an isomorphism of vector spaces. Thus, for finite-dimensional spaces we obtain a natural on-to-one correspondence

$$\left\{ \text{nondegenerate bilinear pairings} \right\}_{\text{on } V} \longleftrightarrow \left\{ \text{isomorphisms} \right\}_{V \to V^\vee}. \quad (42)$$

1.4.6 The second dual space $V^{vv}$

For a general vector space there are no natural nonzero linear transformations $V \to V^\vee$. Each such transformation is equivalent to equipping $V$ with a bilinear pairing (12). There is, however, a natural transformation from $V$ to the dual of its own dual:

$$v \mapsto \hat{v} \in V^{vv} \quad (43)$$

where the linear functional $\hat{v}$ on $V^\vee$ is defined as

$$\hat{v}(\phi) := \phi(v) \quad (\phi \in V^\vee). \quad (44)$$

The value of $\hat{v}$ on $\phi$ is declared to be the value of the functional $\phi$ on the vector $v$. 

16
Exercise 6 Calculate $\hat{v} + \hat{v}'(\phi)$ and $\hat{\alpha}v(\phi)$ and show that the correspondence

$$^\wedge: V \longrightarrow V^\vee$$

(45)
is a linear transformation.

Exercise 7 Show that for any nonzero vector $v$, there exists a linear functional $\phi$ such that $\phi(v) \neq 0$. Deduce from this that $\hat{\phi} = 0$ if and only if $v = 0$. (Hint: consider a basis $B$ in which $b_1 = v$.)

1.4.7
For a finite dimensional vector space we established in Section 1.2.8 that $V$ and its dual $V'$ have the same dimension. In particular, also $V'$ and $V^{\vee\vee}$ have the same dimension. Combining this with the fact that the transformation (45) is, according to Exercise 7, injective, we infer that it is an isomorphism. Thus, a finite-dimensional vector space $V$ naturally identifies with its own second dual.

1.4.8
It follows that the coordinate functionals

$$\psi_1^*, \ldots, \psi_n^*$$

for any basis $\{\psi_1, \ldots, \psi_n\}$ in the dual space $V^\vee$ are necessarily of the form

$$^\wedge b_1, \ldots, ^\wedge b_n,$$

for some basis $B = \{b_1, \ldots, b_n\}$ in $V$.

1.4.9
Let us calculate the value of $\psi_i$ on an arbitrary vector $v$ expressed in basis $B$, cf. (1),

$$\psi_i(v) = \psi(\beta_1 b_1 + \cdots + \beta_n b_n)$$

$$= \beta_1 \psi_i(b_1) + \cdots + \beta_n \psi_i(b_n)$$

$$= \beta_1 ^\wedge b_1(\psi_i) + \cdots + \beta_n ^\wedge b_n(\psi_i)$$

$$= \beta_1 \Psi_i^*(\psi_i) + \cdots + \beta_n \Psi_n^*(\psi_i)$$

$$= \beta_i.$$  

We have, of course, $b_i^*(v) = \beta_i$. Thus, the functionals $\psi_i$ and $b_i^*$ are equal.
1.4.10

We demonstrated that every basis in the dual space $V^\vee$ of a finite-dimensional space consists of the coordinate functionals $B^* = \{b_1^*, \ldots, b_n^*\}$ of a unique basis $B = \{b_1, \ldots, b_n\}$ in $V$.

1.4.11

We also demonstrated that the coordinate functionals $b_1^{**}, \ldots, b_n^{**}$, of the dual basis $B^*$ identify with vectors of the original basis in $V$ via natural identification (43) of $V^{**}$ with $V$,

$$b_1^{**} = \hat{b}_1, \ldots, b_n^{**} = \hat{b}_n. \quad (46)$$
2 Linear transformations

2.1 The adjoint linear transformation

2.1.1 The dual linear transformation

A linear transformation $T: V \rightarrow W$ induces a linear transformation between the dual spaces with the source and the target exchanging places:

$$T^\vee : W^\vee \rightarrow V^\vee, \quad T^\vee (\psi) := \psi \circ T, \quad (\psi \in W^\vee). \quad (47)$$

2.1.2

As you see, $T^\vee$ sends a linear functional $\psi: W \rightarrow \mathbb{R}$ on $W$ to its composition with $T$,

$$V \xrightarrow{T} W \xrightarrow{\psi} \mathbb{R}$$

yielding thus a linear functional on $V$. The dependence on $\psi$ is linear—which just reflects the fact that composition of linear transformations is distributive with respect to addition of linear transformations and commutes with multiplication by scalars,

$$(\psi + \psi') \circ T = \psi \circ T + \psi' \circ T, \quad (\alpha \psi) \circ T = \alpha (\psi \circ T),$$

so $T^\vee$ is a linear transformation from the dual of $W$ to the dual of $V$.

2.1.3 The kernel of $T^\vee$

The kernel of $T^\vee$ is the set of functionals $\psi$ on $W$ such that $\psi \circ T = 0$. These are the functionals that vanish on the range $T(V)$ of $T$:

$$\ker T^\vee = \{ \psi \in W^\vee \mid \ker \psi \supseteq T(V) \}. \quad (48)$$

2.1.4

Suppose that both $V$ and $W$ are equipped with scalar products. We shall denote the scalar product on $V$ by $\langle \, , \rangle$ and the one on $W$ by $\langle \, , \rangle'$. In the diagram

$$\begin{array}{c}
V^\vee \xrightarrow{T^\vee} W^\vee \\
\text{*} & \text{*}' \\
V \xrightarrow{\ast} W
\end{array}$$
the vertical arrows are the isomorphisms associated with the corresponding scalar products
\[ v \mapsto v^* = \langle v, \rangle, \quad w \mapsto w^{*'} = \langle w, \rangle'. \]

By composing \( *' \) with \( T^\vee \) and then with the inverse of \( * \) we obtain the linear transformation \( T^*: W \rightarrow V \) making the following diagram commutative

\[
\begin{array}{ccc}
V^\vee & \xrightarrow{T^\vee} & W^\vee \\
\downarrow{*} & & \downarrow{*'} \\
V & \xrightarrow{T^*} & W
\end{array}
\] (49)

It is called the adjoint of \( T \). Unlike \( T^\vee \), the adjoint is defined on the original spaces, not their duals, but its construction depends on the choice of scalar products on the source and on the target of \( T \).

2.1.5

Note that \( * \circ T^* \) takes a vector \( w \in W \) to
\[ (T^*w)^* = \langle T^*w, \rangle, \]
while \( T^\vee \circ *' \) takes \( w \) to
\[ w^{*'} \circ T = \langle w, T(\rangle). \]

The commutativity of diagram (49) means that the following linear functionals on \( V \) are equal:
\[ \langle T^*w, \rangle = \langle w, T(\rangle, \]
i.e., that, for any vector \( v \in V \), one has
\[ \langle T^*w, v \rangle = \langle w, Tv \rangle'. \] (50)

2.1.6

This last identity is how the adjoint transformation is usually defined but then one has to demonstrate that the desired linear transformation \( T^*: W \rightarrow V \) exists and is unique. The way we define \( T^* \) is free of this inconvenience.
2.1.7 The kernel of $T^*$

We shall record a few immediate consequences of identity (50). In view of the bilinear pairing $\langle \cdot, \cdot \rangle$ being nondegenerate, the vector $T^*w$ is zero if and only if

$$0 = \langle T^*w, v \rangle = \langle w, Tv \rangle', $$

for all $v \in V$, i.e., if and only if $w$ is orthogonal to the range of $T$. Thus,

$$\ker T^* = T(V)\perp. \tag{51}$$

2.1.8 The range of $T^*$

A vector $v \in V$ is orthogonal to the range of $T^*$ if and only $\langle w, Tv \rangle' = 0$ is zero for all $w \in W$. In view of $\langle , \rangle'$ being nondegenerate, this is equivalent to $Tv = 0$. Thus,

$$T^*(W)\perp = \ker T$$

or, equivalently,

$$T^*(W)\perp\perp = \ker T\perp. \tag{52}$$

In Section 4.1.4 we shall establish that for finite-dimensional subspaces $U \subseteq V$ one has $U\perp\perp = U$, hence

$$T^*(W) = \ker T\perp \tag{53}$$

since we assume both $V$ and $W$ to be finite-dimensional.

2.1.9

For any vector $v \in V$ in a vector space with a scalar product we defined in Section 1.4.1 the associated linear functional $v^*$. In fact, $v^*$ is the adjoint (!) of the linear transformation

$$\mathbf{R} \rightarrow V,$$

sending a real number $a$ to $av$. Isn’t this remarkable? Mathematics is full of such beautiful miracles, one only needs patience and readiness to discover them (the lazy ones are deprived of many joys, including spectacular joys of such discoveries).

Note that vectors in $V$ are here naturally identified with linear transformations from $\mathbf{R}$ to $V$. 

21
2.2 Linear transformations and bases

2.2.1

Let $T: V \to W$ be a linear transformation. Suppose that $V$ is equipped with a basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ while $W$ is equipped with a basis $\mathcal{C} = \{c_1, \ldots, c_m\}$. The image of each basis vector $b_j$ under $T$ is a linear combination of basis vectors in $W$,

$$Tb_j = \alpha_{1j}c_1 + \cdots + \alpha_{mj}c_m. \quad (54)$$

If $v$ is an arbitrary vector (1) in $V$, then

$$Tv = T(\beta_1b_1 + \cdots + \beta_nb_n)$$
$$= \beta_1 Tb_1 + \cdots + \beta_n Tb_n$$
$$= \beta_1(\alpha_{11}c_1 + \cdots + \alpha_{m1}c_m) + \cdots + \beta_n(\alpha_{1n}c_1 + \cdots + \alpha_{mn}c_m)$$
$$= \sum_{1 \leq j \leq n} \sum_{1 \leq i \leq m} \beta_j \alpha_{ij}c_i$$
$$= \sum_{1 \leq i \leq m} \left( \sum_{1 \leq j \leq n} \alpha_{ij} \beta_j \right) c_i$$

In other words, the column of coefficients representing $Tv$ in basis $\mathcal{C}$, equals $A$ times the column of coefficients representing $v$ in basis $\mathcal{B}$,

$$[Tv]_\mathcal{C} = A[v]_\mathcal{B}. \quad (55)$$

Here $A$ denotes the $m \times n$ matrix with columns

$$[Tb_1]_\mathcal{C}, \ldots, [Tb_n]_\mathcal{C}.\]$$

2.2.2 A commutative diagram of linear transformations

Identity (55) expresses the fact that the following diagram of linear transformations commutes

$$\begin{array}{ccc}
\mathbb{R}^n & \xrightarrow{L_A} & \mathbb{R}^m \\
\downarrow & & \downarrow \\
V & \xrightarrow{T} & W
\end{array}$$

which means that the composition of the top arrow with the left arrow produces the same transformation as the composition of the right arrow
with the bottom arrow. Here \( L_A \) denotes left multiplication by matrix \( A \) and the vertical arrows are the coordinatization isomorphisms of \( V \) with \( \mathbb{R}^n \) and \( W \) with \( \mathbb{R}^m \), induced by basis \( \mathcal{B} \) in \( V \) and basis \( \mathcal{C} \) in \( W \), respectively.

### 2.2.3

The matrix \( A \) depends on \( T \) and on the choice of bases in both the source and the target spaces of \( T \). When necessary, we shall indicate this dependence by employing the notation \( A_T(\mathcal{C}, \mathcal{B}) \).

### 2.2.4

If \( S: W \to X \) is another linear transformation and \( \mathcal{D} = \{d_1, \ldots, d_l\} \) is a basis in \( X \), then the commutativity of the diagram

\[
\begin{array}{ccc}
\mathbb{R}^n & \xrightarrow{L_{A_T(\mathcal{C}, \mathcal{B})}} & \mathbb{R}^m & \xrightarrow{L_{A_S(\mathcal{D}, \mathcal{C})}} & \mathbb{R}^l \\
\| \|_B & \downarrow & \| \|_C & \downarrow & \| \|_D \\
V & \xrightarrow{T} & W & \xrightarrow{S} & X
\end{array}
\]

means that

\[
A_S(\mathcal{D}, \mathcal{C})A_T(\mathcal{C}, \mathcal{B}) = A_{S \circ T}(\mathcal{D}, \mathcal{B}).
\]

### 2.2.5

Thus, we obtain a one-to-one correspondence

\[
\left\{ \text{linear transformations } T: V \to W \right\} \longleftrightarrow \left\{ m \times n \text{ matrices} \right\}
\]

where \( m \) is the dimension of the target space and \( n \) is the dimension of the source space.

---

\(^1\)Warning to my Math 54 students: in my lectures I have been using a slightly different notation, \( A_T(\mathcal{B}, \mathcal{C}) \).
2.2.6

This correspondence is not natural: it depends on the chosen bases in $V$ and $W$. It is, however, compatible with the operations on the spaces of linear transformations and the corresponding spaces of matrices, see (58) and

$$A_{T+T'}(\mathcal{C}, \mathcal{B}) = A_T(\mathcal{C}, \mathcal{B}) + A_{T'}(\mathcal{C}, \mathcal{B}), \quad A_{\alpha T}(\mathcal{C}, \mathcal{B}) = \alpha A_T(\mathcal{C}, \mathcal{B}). \quad (60)$$

The pair of above identities means that the correspondence

$$T \mapsto A_T(\mathcal{C}, \mathcal{B})$$

establishes an isomorphism between the vector space $\text{Lin}(V, W)$ of linear transformations from $V$ to $W$ and the vector space $\text{Mat}_{mn}(\mathbb{R})$ of $m \times n$ matrices.

2.2.7 The case of $T : \mathbb{R}^n \to \mathbb{R}^m$

In the special case when $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ and the bases chosen are the standard bases, the coordinatization isomorphisms are the identity transformations and the commutative diagram becomes

$$\begin{array}{ccc}
\mathbb{R}^n & \overset{L_A}{\longrightarrow} & \mathbb{R}^m \\
\text{id} & & \text{id} \\
\mathbb{R}^n & \overset{T}{\longrightarrow} & \mathbb{R}^m
\end{array} \quad (61)$$

meaning that every linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ is of the form $L_A$ for a unique $m \times n$ matrix. The columns of $A$ are the images of the standard basis vectors $e_1, \ldots, e_n \in \mathbb{R}^n$ under $T$,

$$Te_1, \ldots, Te_n.$$
2.2.8 The effect of a change of basis on the matrix $A_T(\mathcal{C}, \mathcal{B})$

Given another pair of bases $\mathcal{B}'$ and $\mathcal{C}'$ in $V$ and $W$, respectively, we obtain a commutative diagram

\[
\begin{array}{c}
\mathbb{R}^n \xrightarrow{L_A} \mathbb{R}^m \\
\begin{array}{c|c}
\mathcal{B} & \mathcal{C} \\
\hline
V & W \\
\hline
\mathcal{B}' & \mathcal{C}' \\
\end{array}
\end{array}
\]

and, therefore, also the diagram

\[
\begin{array}{c}
\mathbb{R}^n \xrightarrow{L_{A'}} \mathbb{R}^m \\
\begin{array}{c|c}
\mathcal{B} & \mathcal{C} \\
\hline
V & W \\
\hline
\mathcal{B}' & \mathcal{C}' \\
\end{array}
\end{array}
\]

2.2.9 The change-of-coordinates matrices

The transformation $[ \cdot ]_{\mathcal{B}} \circ [ \cdot ]_{\mathcal{B}'}^{-1}$ sends the standard vectors $e_1, \ldots, e_n$ to vectors $b'_1, \ldots, b'_n$ and then represents them in basis $\mathcal{B}$ as columns of numbers. It acts, therefore, as multiplication by the $n \times n$ matrix with columns

$[b'_1]_{\mathcal{B}}, \ldots, [b'_n]_{\mathcal{B}}$.  

$Lay$ denotes this matrix $P_{\mathcal{B}' \leftarrow \mathcal{B}}$ and calls it the change-of-coordinates matrix from $B'$ to $B$.

2.2.10

The inverse of the change-of-coordinates matrix $P_{\mathcal{B}' \leftarrow \mathcal{B}}$ is, by definition, $P_{\mathcal{B} \leftarrow \mathcal{B}'}$.
2.2.11 The change-of-coordinates formula

The commutativity of diagram (63) translates into the formula

\[ A_T(c', b') = P_{c' \to c} A_T(c, b) P_{b \to b'}. \]  \hspace{1cm} (65)

2.3 The matrix of the dual transformation

2.3.1

Every linear functional \( \psi : \mathbb{R}^m \to \mathbb{R} \) is of the form

\[ \psi = L_\alpha \]

where \( \alpha = [\alpha_1 \ldots \alpha_m] \) is a \( 1 \times m \) matrix, i.e., a row of \( m \) numbers. Therefore the dual of the transformation

\[ T = L_A : \mathbb{R}^n \to \mathbb{R}^m \]

sends \( \psi \) to the composite \( \psi \circ T = L_\alpha \circ L_A = L_{\alpha A} \).

2.3.2

If we identify \( (\mathbb{R}^m)\vee \) and \( (\mathbb{R}^n)\vee \) with the spaces of row vectors \( \text{Mat}_{1m}(\mathbb{R}) \) and, respectively, \( \text{Mat}_{1n}(\mathbb{R}) \), then \( (L_A)\vee \) becomes right multiplication by matrix \( A \),

\[ (L_A)\vee (\alpha) = \alpha A. \] \hspace{1cm} (66)

2.3.3

The coordinatization of the dual space \( (\mathbb{R}^n)\vee \) is the composite of the identification

\( (\mathbb{R}^n)\vee \simeq \text{Mat}_{1n}(\mathbb{R}) \)

and transposition

\[ \tau : \text{Mat}_{1n}(\mathbb{R}) \to \text{Mat}_{n1}(\mathbb{R}) = \mathbb{R}^n. \]

Therefore, the coordinatization of \( (\mathbb{R}^m)\vee \) and \( (\mathbb{R}^n)\vee \) identifies \( (L_A)\vee \) with left multiplication by the transpose matrix \( A^\tau \),

\[ y \mapsto A^\tau y \quad (y \in \mathbb{R}^m). \] \hspace{1cm} (67)
The inverse to the coordinatization of \((\mathbb{R}^n)^\vee\) is the isomorphism

\[ * : \mathbb{R}^n \rightarrow (\mathbb{R}^n)^\vee \]  

associated with the dot product on \(\mathbb{R}^n\).

Let \(V\) be a vector space with a basis \(\mathcal{B}\). The composite of (68) and the dual of the coordinatization isomorphism \([\ ]_{\mathcal{B}}\),

\[ ([\ ]_{\mathcal{B}})^\vee : (\mathbb{R}^n)^\vee \rightarrow V^\vee, \]

sends the standard basis vectors \(e_i\) in \(\mathbb{R}^n\) to the basis vectors \(b_i\) in \(V\). This means that the inverse transformation is the coordinatization of \(V^\vee\) in the dual basis \(\mathcal{B}^*\),

\[ [\ ]_{\mathcal{B}^*} = \left(([\ ]_{\mathcal{B}})^\vee \circ (\ )^*\right)^{-1}. \]  

In particular, if \(T : V \rightarrow W\) is a linear transformation with matrix \(A\) in bases \(\mathcal{B}\) in \(V\) and \(\mathcal{C}\) in \(W\), then the matrix of its dual

\[ T^\vee : W^\vee \rightarrow V^\vee \]

is the transpose matrix \(A^\tau\),

\[ A_{T^\vee} (\mathcal{B}^*, \mathcal{C}^*) = A_T (\mathcal{C}, \mathcal{B})^\tau. \]  

The rank of a linear transformation

The rank of \(T : V \rightarrow W\) is the dimension of its range, i.e., the vector subspace \(T(V)\) of \(W\) consisting of the values of \(T\).

Exercise 8 Suppose that vectors \(Tv_1, \ldots, Tv_l\) are linearly independent. Show that vectors \(v_1, \ldots, v_l\) are linearly independent. (Hint: show that if \(v_1, \ldots, v_l\) are linearly dependent, then \(Tv_1, \ldots, Tv_l\) are linearly dependent.)

27
2.4.2

The matrix representation of a linear transformation, as described by identity (55), or by the commutative diagram (56), can be also expressed in the following form

\[ T = \sum_{1 \leq i \leq m} \sum_{1 \leq j \leq n} a_{ij} c_i b_j^*, \]  
(71)

where \( b_1^*, \ldots, b_n^* \) are the coordinate functionals on \( V \).

2.4.3

Each term \( c_i b_i^* \) is a linear transformation whose range is the one-dimensional subspace of \( W \) spanned by \( c_i \). Its rank is thus 1 and the same is true when we multiply \( c_i b_i^* \) by a nonzero coefficient \( a_{ij} \). Thus, the matrix representation of a linear transformation represents it as the sum of rank 1 transformations whose number equals the number of nonzero entries in matrix \( A \). This number does not exceed \( mn \) and often is equal to it.

2.4.4

Let us note that the identity operator \( \text{id}_V \) on a vector space \( V \) of dimension \( n \) has obviously rank \( n \) and it also has the following representation as the sum of \( n \) rank 1 operators,

\[ \text{id}_V = b_1^* b_1 + \cdots + b_n^* b_n, \]  
(72)

where \( \{b_1, \ldots, b_n\} \) is an arbitrary basis of \( V \). This is nothing more than observing that the identity operator is represented by the identity \( n \times n \) matrix

\[
\begin{bmatrix}
1 \\
\vdots \\
1
\end{bmatrix}
\]

in every basis.

2.4.5

If a linear transformation \( T: V \to W \) can be expressed as the sum of \( d \) rank 1 transformations, then the rank of \( T \) does not exceed \( d \) because the range of \( T \) is spanned by \( d \) nonzero vectors belonging to the one-dimensional ranges of those rank 1 transformations.
This means that a transformation of rank \( r \) cannot be represented as the sum of fewer than \( r \) rank 1 transformations. We shall now show that, in fact, any linear transformation of rank \( r \) can be represented by exactly \( r \) transformations of rank 1.

### 2.4.6

Indeed, let \( \{c_1, \ldots, c_r\} \) be any basis in the range of \( T \). Let

\[
\tau_1 := T^\ast(c_1^\ast), \quad \ldots, \quad \tau_r := T^\ast(c_r^\ast).
\]

Then

\[
(c_1 \tau_1 + \cdots + c_r \tau_r)(v) = c_1 \tau_1(v) + \cdots + c_r \tau_r(v)
\]

\[
= c_1 c_1^\ast(Tv) + \cdots + c_r c_r^\ast(Tv)
\]

\[
= (c_1 c_1^\ast + \cdots + c_r c_r^\ast)(Tv)
\]

\[
= \text{id}_{\text{Range of } T}(Tv) = Tv.
\]

We demonstrated that \( T \) has the following representation as the sum of \( r \) rank 1 linear transformations

\[
T = c_1 \tau_1 + \cdots + c_r \tau_r. \quad (73)
\]

### 2.4.7

We shall now demonstrate that for any linear transformation \( T \) it is possible to find such bases \( B \) on \( V \) and \( C \) on \( W \), that the functionals \( \tau_1, \ldots, \tau_r \) are the coordinate functionals of \( B \),

\[
\tau_1 = b_1^\ast, \quad \ldots, \quad \tau_r = b_r^\ast.
\]

### 2.4.8

Indeed, by performing row and column operations it is possible to represent any \( m \times n \) matrix \( A \) of rank \( r \) as the product

\[
A = PJQ \quad (74)
\]

where \( P \) is an invertible \( m \times m \) matrix, \( Q \) is an invertible \( n \times n \) matrix, and \( J \) is the \( m \times n \) matrix having the diagonal \( r \times r \) identity matrix in its
left top corner and everywhere else zeros,

\[
J = \begin{bmatrix}
1 & \cdots & 1 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 1 \\
\end{bmatrix}
\] (75)

Matrix \( J \) is the reduced column echelon form of the reduced row echelon form of \( A \).

2.4.9

If \( A \) is the matrix of \( T \) in some bases \( \mathcal{B}' \) and \( \mathcal{C}' \), then, according to the change-of-coordinates formula (65), \( J \) will become the matrix of \( T \) in the bases \( \mathcal{B} \) and \( \mathcal{C} \) such that

\[ P = P_{\mathcal{C}' \rightarrow \mathcal{C}} \quad \text{and} \quad Q = P_{\mathcal{B} \rightarrow \mathcal{B}'} \].

Saying that \( A_T(\mathcal{C}, \mathcal{B}) = J \) means that \( T \) has the following representation as the sum of \( r \) rank 1 transformations:

\[ T = c_1 b_1^* + \cdots + c_r b_r^* \] (76)

and this is the simplest such representation of \( T \).

2.5 The row rank versus the column rank of a matrix

2.5.1

Given an \( m \times n \) matrix \( A \), its row rank is the dimension of the vector space spanned by the rows of \( A \)

\[ \text{Row } A := \text{lin span}\{A^1, \ldots, A^m\} \] (77)

while the column rank is the dimension of the vector space spanned by the columns of \( A \)

\[ \text{Col } A := \text{lin span}\{A_1, \ldots, A_n\} \] (78)
Note that Row $A$ is the range of right multiplication by $A$,

$$R_A: \text{Mat}_{1m}(\mathbb{R}) \longrightarrow \text{Mat}_{1n}(\mathbb{R}),$$

whereas Col $A$ is the range of left multiplication by $A$,

$$L_A: \text{Mat}_{n1}(\mathbb{R}) \longrightarrow \text{Mat}_{m1}(\mathbb{R}).$$

2.5.2

For any linear transformation $T: V \rightarrow W$ and any isomorphisms $S: W \rightarrow W'$ and $U: V' \rightarrow V$, the range of $T$ coincides with the range of $T \circ U$ and is isomorphic to the range of $S \circ T$ (the isomorphism is provided by $S$). It follows that

$$\dim \text{Range} (S \circ T \circ U) = \dim \text{Range} T. \quad (79)$$

2.5.3

By applying this to $T$ being either $L_A$ or $R_A$, we obtain that both the row and the column ranks remain constant when one multiplies the matrix on either side by an invertible matrix. Thus, the row and the column ranks of $A = PJQ$ equal the row and, respectively, the column ranks of matrix $J$, cf. (74)–(75). The two ranks are equal for matrix $J$, hence they are equal for the original matrix $A$. 

31
3 Linear operators on $V$

3.1 Projections

3.1.1
Linear transformations $T: V \to V$ are usually referred to as linear operators on a vector space $V$.

3.1.2
An operator $P$ on a vector space $V$ is said to be a projection onto a subspace $W$ along a subspace $W'$ if its range is $W$, its kernel is $W'$, and
\[
P(v) = v \quad \text{for all} \quad v \in W.
\] (80)

3.1.3 Idempotent operators
An operator $T$ is said to be idempotent if
\[
T \circ T = T.
\] (81)

3.1.4
Any projection is idempotent. Indeed, $P(v)$ belongs to $W$, for any vector $v$, and therefore $P(P(v)) = P(v)$, in view of (80).

3.1.5
For an idempotent operator $T$, denote by $W'$ its kernel and by $W$ its range. Any vector $w$ in $W$ is of the form $w = T(v)$ for some $v \in V$. Hence
\[
T(w) = T(T(v)) = T(v) = w,
\]

i.e., $T$ is a projection onto its range along its kernel.

3.1.6 The complementary projection
If $P$ is a projection, then $P' := \text{id}_V - P$ is an idempotent operator too:
\[
P' \circ P' = (\text{id}_V - P) \circ (\text{id}_V - P)
= \text{id}_V - 2P + P \circ P
= \text{id}_V - 2P + P = P'.
\]
Moreover,
\[ P'(v) = 0 \quad \text{if and only if} \quad P(v) = v \]
and
\[ P'(v) = v \quad \text{if and only if} \quad P(v) = 0, \]
so \( P' \) is a projection onto the kernel of \( P \) along the range of \( P \).

3.1.7

Note that
\[ P \circ P' = 0 = P' \circ P. \] (82)

It is customary to say about two projections satisfying (82) that they are \textit{orthogonal} to each other.

3.1.8

The operator \( P' \) is referred to as the \textbf{complementary projection}. Its range is \( W' \) and its kernel is \( W \).

3.1.9  \textbf{The associated direct sum decomposition of} \( V \)

Since
\[ \text{id}_V = P + P', \]
any vector \( v \in V \) is the sum
\[ v = w + w' \] (83)
for some \( w \in W \) and \( w' \in W' \) (namely, \( w = P(v) \) and \( w' = P'(v) \)), and such representation is unique. Indeed, if
\[ v = w_1 + w'_1 \]
is another representation, then
\[ w_1 = P(v) = w \quad \text{and} \quad w'_1 = P'(v) = w'. \]

3.1.10

When any vector in a vector space \( V \) can be represented as the sum of vectors (83) from two subspaces \( W \) and \( W' \), and that representation is unique, we say that \( V \) is the \textbf{direct sum} of subspaces \( W \) and \( W' \). Note that the intersection of \( W \) with \( W' \) contains only the zero vector.
More generally, when any vector in a vector space $V$ can be represented as the sum of vectors

$$v = w_1 + \cdots + w_l$$  \hfill (84)

from subspaces $W_1, \ldots, W_l$ of $V$, and the representation (84) is unique, then we say that $V$ is the **direct sum** of subspaces $W_1, \ldots, W_l$. In this case, we shall refer to the vectors $W_1 \in W_1, \ldots, w_l \in W_l$ as the **component vectors** of vector $v$.

### 3.1.12 The associated projections

Let $W_i'$ be the set of vectors $v$ in $V$ whose $i$-th component vector is zero. It is a vector subspace such that $V$. If we rewrite (84) as

$$v = w_i + w_i'$$

where $w_i'$ is the sum of all vectors but $w_i$ on the right hand side of (84), then we see that $V$ is the direct sum of $W_i$ and $W_i'$.

Let $P_i$ be the corresponding projection onto $W_i$ with kernel $W_i'$. Since the range of $P_j$ is contained in kernel of $P_i$ when $i \neq j$, the associated projections are orthogonal to each other:

$$P_i \circ P_j = 0 \quad \text{if} \quad i \neq j.$$

The fact that every vector in $V$ has a representation (84) means that

$$\text{id}_V = P_1 + \cdots + P_l.$$  \hfill (85)

### 3.1.13 Bases compatible with a direct sum decomposition

Choosing a basis $\mathcal{B}_i = \{b_1^{(i)}, \ldots, b_{n_i}^{(i)}\}$ in each $W_i$, allows to represent each component vector $w_i$ as a unique linear combination

$$w_i = \alpha_1^{(i)} b_1^{(i)} + \cdots + \alpha_{n_i}^{(i)} b_{n_i}^{(i)}$$  \hfill (86)

of vectors in $\mathcal{B}_i$ where $n_i$ denotes the dimension of $W_i$. By substituting these linear combinations into (84), we obtain a representation of any vector in $V$ as a linear combination of the set $\mathcal{B}$ assembled by taking the union of the component bases $\mathcal{B}_1, \ldots, \mathcal{B}_l$. Such a representation of $v$ as a linear combination of vectors from $\mathcal{B}$ is unique in view of the uniqueness of the representation (84) and of the representations (86). In particular, $\mathcal{B}$ is a basis of $V$. 

34
In particular, the dimension of $V$ is the sum of the dimensions of the component subspaces

$$n = n_1 + \cdots + n_l.$$  \hfill (87)

Any basis $B = \{b_1, \ldots, b_n\}$ in an arbitrary vector space is compatible with several direct sum decompositions of $V$. For example, let us partition $\{1, \ldots, n\}$ into two disjoint subsets $I$ and $J$ and let $V_I$ and, respectively, $V_J$, be the subspaces spanned by basis vectors $b_i$ where $i$ belongs to $I$ or, respectively, to $J$. Then $V$ is the direct sum of $V_I$ and $V_J$.

Another example: let $V_i$ be the one-dimensional subspace spanned by the single vector $b_i$. Then $V$ is the direct sum of one-dimensional subspaces $V_1, \ldots, V_n$.

3.2 Case study: a linear transformation $T: V \to W$

3.2.1

Given a basis $\{c_1, \ldots, c_r\}$ in the range of a linear transformation $T: V \to W$, let us choose vectors $b_1, \ldots, b_r$ in $V$ such that

$$Tb_1 = c_1, \ldots, Tb_r = c_r.$$  

They are linearly independent. Indeed, if

$$\alpha_1 b_1 + \cdots + \alpha_r b_r = 0,$$

then

$$\alpha_1 c_1 + \cdots + \alpha_r c_r = T(\alpha_1 b_1 + \cdots + \alpha_r b_r) = 0$$

and the linear independence of $c_1, \ldots, c_r$ implies that all $\alpha_i$ are zero. This argument, in fact, shows that no nonzero linear combination of vectors $b_1, \cdots, b_r$ belongs to the kernel of $T$.  

35
3.2.2

Let us denote the linear span of vectors $b_1, \ldots, b_r$ by $\tilde{V}$ and the kernel of $T$ by $V_0$. Given a vector $v \in V$, let us represent $Tv$ in basis $\{c_1, \ldots, c_r\}$,
\[ Tv = \alpha_1 c_1 + \cdots + \alpha_r c_r. \]
Then $\tilde{v} := \alpha_1 b_1 + \cdots + \alpha_r b_r$ belongs to $\tilde{V}$ and
\[ T(v - \tilde{v}) = (\alpha_1 c_1 + \cdots + \alpha_r c_r) - T(\alpha_1 b_1 + \cdots + \alpha_r b_r) = 0 \]
means that $v - \tilde{v}$ belongs to $V_0$.

3.2.3

Thus, $V$ is the direct sum of $\tilde{V}$ and $V_0$. In particular, we demonstrated that
\[ \dim V = \dim T(V) + \dim \ker T \quad (88) \]
because $\tilde{V}$ is spanned by $r$ linearly independent vectors $b_1, \ldots, b_r$ and $r$ is the dimension of the range of $T$.

We also showed that
- for any basis $\{c_1, \ldots, c_r\}$ in the range of $T$, there exists a basis $\{b_1, \ldots, b_n\}$ in $V$ such that (76) holds.

3.3 Linear operators on a space with a chosen basis

3.3.1

Given a basis $\mathcal{B}$ on $V$, the matrix $A_T(\mathcal{B}, \mathcal{B})$ will be referred to as the matrix of $T$ in basis $\mathcal{B}$. We shall denote it $A_T(\mathcal{B})$.

3.3.2

For two operators $S$ and $T$ and a number $a$, one has
\[ A_{S+T}(\mathcal{B}) = A_S(\mathcal{B}) + A_T(\mathcal{B}), \quad A_{aT}(\mathcal{B}) = aA_T(\mathcal{B}), \quad (89) \]
and
\[ A_{S\circ T}(\mathcal{B}) = A_S(\mathcal{B})A_T(\mathcal{B}). \quad (90) \]
These are special cases of identities (60) and (58).
3.3.3 Polynomial functions of an operator

Let

\[ p(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_d t^d \]  

be a polynomial. By substituting an operator \( T \) in place of the symbol \( t \), we obtain an operator \( p(T) \) on \( V \). Identities (60)–(90) imply that the matrix of \( p(T) \) is that polynomial applied to the matrix of \( T \),

\[ A_{p(T)}(B) = p(A_T(B)). \]  

3.4 Invariant subspaces of a linear operator

A subspace \( W \) of \( V \) is said to be invariant under an operator \( T \) if

\[ T(W) \subseteq W, \]

i.e., if \( Tw \) belongs to \( W \) for any vector \( w \in W \). The intersection

\[ W_1 \cap \cdots \cap W_l \]

and the span

\[ W_1 + \cdots + W_l \]

of any family \( W_1, \cdots, W_l \) of invariant subspaces is invariant. The whole \( V \) is the largest invariant subspace of \( V \) and the zero space \( \{0\} \), containing only the zero vector, is the smallest.

3.4.1 Linear operators on one-dimensional spaces

Every operator on a one-dimensional space \( V \) is of the form

\[ v \mapsto \alpha v \quad (v \in V) \]

for a unique real number \( \alpha \). Indeed, if \( v \neq 0 \), then \( v \) spans \( V \) and therefore \( Tv \) is a multiple \( \alpha v \) for some number \( \alpha \). The coefficient \( \alpha \) does not depend on \( v \) since for \( v' = \alpha' v \), one has

\[ T(v') = T(\alpha' v) = \alpha' T(v) = \alpha' \alpha v = \alpha \alpha' v = \alpha v'. \]

One-dimensional spaces have no nontrivial subspaces so, obviously, operators on one-dimensional spaces have no nontrivial invariant subspaces.
3.5 An example: rotation in $\mathbb{R}^2$

The rotation of the two-dimensional space $\mathbb{R}^2$

$$T = L_A, \quad \text{where} \quad A = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}, \quad (93)$$

has no nontrivial invariant subspaces.

3.6 One-dimensional invariant subspaces and eigenvalues

Restriction of $T$ to any of its invariant subspaces $W$ induces an operator on $W$. If $W$ is one-dimensional, $T$ acts on $W$ as multiplication by a certain number $\lambda$. The vectors of one-dimensional invariant subspaces are called eigenvectors of $T$ with eigenvalue $\lambda$. The set of all eigenvectors with $\lambda$ as their eigenvalue forms an invariant subspace that is called the eigensubspace of $T$ corresponding to eigenvalue $\lambda$. It will be denoted $V_\lambda$.

3.7

By definition, $V_\lambda$ is the largest subspace of $V$ on which $T$ acts as multiplication by $\lambda$.

3.8 The subspace spanned by all eigenvectors

Denote by $V'$ the subspace of $V$ spanned by all eigenvectors of $T$. We shall assume in what follows that $V$ is finite-dimensional. In this case, there are finitely many distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and

$$V' = V_{\lambda_1} + \cdots + V_{\lambda_k}. \quad (94)$$

We shall show that any vector $v$ in $V'$ has a unique representation

$$v = v_1 + \cdots + v_k \quad (v_1 \in V_{\lambda_1}, \ldots, v_k \in V_{\lambda_k}).$$

Indeed, if

$$v = v'_1 + \cdots + v'_k \quad (v'_1 \in V_{\lambda_1}, \ldots, v'_k \in V_{\lambda_k})$$

is another representation, then

$$(v_1 - v'_1) + \cdots + (v_k - v'_k) = 0 \quad (95)$$
By applying $T^i$ to both sides of (95) we obtain the system of equalities

$$
(v_1 - v'_1) + \cdots + (v_k - v'_k) = 0
$$
$$
\lambda_1(v_1 - v'_1) + \cdots + \lambda_k(v_k - v'_k) = 0
$$
$$
\lambda_1^2(v_1 - v'_1) + \cdots + \lambda_k^2(v_k - v'_k) = 0
$$

(96)

Consider the vector space $V^k$ whose elements are columns of $k$ vectors in $V$. The first $k$ equalities in (96) together express the fact that the column of vectors

$$
\begin{bmatrix}
  v_1 - v'_1 \\
  \vdots \\
  v_k - v'_k
\end{bmatrix}
$$

(97)

is annihilated by the operator on $V^k$ that multiplies a column of vectors by the square matrix

$$
\begin{bmatrix}
  1 & \cdots & 1 \\
  \lambda_1 & \cdots & \lambda_k \\
  \lambda_1^2 & \cdots & \lambda_k^2 \\
  \vdots & \cdots & \vdots \\
  \lambda_1^{k-1} & \cdots & \lambda_k^{k-1}
\end{bmatrix}
$$

(98)

This matrix is known as the *Vandermonde matrix*. Its determinant equals

$$
\prod_{1 \leq i < j \leq k} (\lambda_i - \lambda_j) \neq 0. 
$$

(99)

In particular, the operator in question is invertible, hence the column of vectors (97) is the zero vector in $V^k$, i.e., all its components are zero.

3.8.1

We demonstrated that the subspace $V'$ spanned by all the eigenvectors of $T$ is the direct sum of the eigenspaces $V_{\lambda_1}, \ldots, V_{\lambda_k}$.

3.8.2

Restriction of $T$ to $V'$ has a very simple matrix $A$ in any basis compatible with the decomposition of $V'$ into the direct sum of the eigenspaces. It is the diagonal matrix having on its diagonal $n_1$ times $\lambda_1$, then $n_2$ times $\lambda_2$, and so on. Here $n_j = \dim V_{\lambda_j}$ denotes the dimension of the space of eigenvectors with eigenvalue $\lambda_j$. 
3.8.3 Diagonalizable operators

We say that an operator $T$ is **diagonalizable** if there is a basis $\mathcal{B}$ on $V$ such that the associated matrix $A_T(\mathcal{B})$ is diagonal. Any such basis consists of nonzero eigenvectors of $T$. It follows that $T$ is diagonalizable if and only if $V = V'$, i.e., if $V$ is spanned by the eigenvectors of $T$. The latter happens precisely when

$$ \dim V = \dim V_{\lambda_1} + \cdots + \dim V_{\lambda_k}. $$

(100)

3.8.4

Since each $\dim V_{\lambda_j} \geq 1$, we infer that any operator having $n = \dim V$ distinct eigenvalues is automatically diagonalizable. In this case each eigenspace is one-dimensional, i.e., each nonzero eigenvector is unique up to a multiple.

**Exercise 9** Describe the eigenvalues and eigenspaces of a projection $P$ and explain why every projection is diagonalizable.
4 Linear operators on a space with a chosen scalar product

4.1 Orthogonal projections

4.1.1 Suppose that $W$ is a subspace equipped with a basis $\{c_1, \ldots, c_r\}$ orthogonal with respect to a scalar product $\langle \ , \ \rangle$ on $V$. This means that

$$c_i \perp c_j \quad (1 \leq i \neq j \leq r).$$

(101)

Consider the operator

$$P := \frac{1}{\langle c_1, c_1 \rangle} c_1^* c_1 + \cdots + \frac{1}{\langle c_r, c_r \rangle} c_r^* c_r$$

(102)

where the functionals $c_i^* \in V^*$ are defined in terms of the scalar product $\langle \ , \ \rangle$, as in Section 1.4.1,

$$c_i^*(v) := \langle c_i, v \rangle \quad (v \in V).$$

4.1.2 The range of $P$ is contained in $W$. In view of (101),

$$Pc_j = \sum_{1 \leq i \leq r} \frac{\langle c_i, c_j \rangle}{\langle c_i, c_i \rangle} c_i = \frac{\langle c_j, c_j \rangle}{\langle c_j, c_j \rangle} c_j = c_j.$$

It follows that $Pv = v$ for each $v \in W$. In particular, $W$ is the range of $P$ and $P$ is a projection onto $W$. The kernel of $P$ consists of vectors

$$v - Pv$$

where $v$ is any vector of $V$.

Exercise 10 Show that $(v - Pv) \perp c_j$ for each $1 \leq j \leq r$.

Any vector orthogonal to $W$ is annihilated by $P$, as follows immediately from the definition of $P$. Thus, the kernel of $P$ coincides with $W^\perp$. 

41
4.1.3 The orthogonal complement of a subspace

We demonstrated that $V$ is the direct sum of $W$ and $W^\perp$. The latter is called the **orthogonal complement** of $W$ (in $V$). The operator defined by (102) is the projection onto $W$ along $W^\perp$. Its definition depends on the choice of an orthogonal basis in $W$ yet the operator itself—does not. We call it the **orthogonal projection onto** $W$.

4.1.4

If $V$ is finite-dimensional, then $W^\perp$ has a finite basis. By applying the above argument to $W^\perp$ instead of $W$, we deduce that $V$ is the direct sum of $W^\perp\perp$ and $W^\perp$. Thus we have two direct sum decompositions of $V$: as the sum of $W$ and $W^\perp$, and as the sum of $W^\perp\perp$. But $W \subseteq W^\perp\perp$, see Exercise 3. The uniqueness of representation of

$$v = w'' + w'$$

where $w'' \in W^\perp\perp$ and $w' \in W^\perp$, then implies that every $w'' \in W^\perp\perp$ belongs to $W$. This means that

$$W^\perp\perp = W.$$  \hfill (103)

4.2 Gram-Schmidt orthogonalization

4.2.1

Any collection of nonzero vectors $w_1, \ldots, w_r$ spanning a vector subspace $W$ can be transformed into an orthogonal basis of $W$. Consider the nondecreasing sequence of vector subspaces

$$W_1 \subseteq W_2 \subseteq \cdots \subseteq W_r = W,$$  \hfill (104)

where $W_i := \text{linspan}\{w_1, \ldots, w_i\}$. Let $P_i$ denote the orthogonal projection onto $W_i$. Then

$$u_1 := w_1, \quad u_2 := w_2 - P_1 w_2, \quad \ldots, \quad u_r := w_r - P_{r-1} w_r,$$  \hfill (105)

defines a sequence of mutually orthogonal vectors such that

$$\text{linspan}\{u_1, \ldots, u_i\} = \text{linspan}\{w_1, \ldots, w_i\} = W_i.$$  

Note that each $u_i$ is the orthogonal projection of $w_i$ onto $W_{i-1}^\perp$ and it is zero precisely when $W_i = W_{i-1}$.  

42
4.2.2

In order to calculate \( u_i \) we need the projection operator \( P_{i-1} \) but that one is provided by the previously calculated vectors \( u_1, \ldots, u_{i-1} \) with help of the formula

\[
P_{i-1} = \frac{1}{\langle u_1, u_1 \rangle} u_1 u_1^* + \cdots + \frac{1}{\langle u_{i-1}, u_{i-1} \rangle} u_{i-1} u_{i-1}^*
\]

(106)

(in this formula we skip the terms corresponding to those \( u_i \) that are zero). Thus, the process of calculating the sequence \( u_1, \ldots, u_r \) becomes a simple algorithm for which we only need the original sequence of vectors \( w_1, \ldots, w_r \). By removing those \( u_i \) that are zero, we obtain an orthogonal basis for \( W \).

The algorithm is known as the **Gram-Schmidt orthogonalization** of any sequence of nonzero vectors. It produces, in particular, from any basis of \( V \), an orthogonal basis. If the original basis is orthogonal, its Gram-Schmidt orthogonalization coincides with it.

### 4.3 Adjoint operators

#### 4.3.1

Given operators \( S \) and \( T \) on \( V \), let us consider the linear pairings that associate to a pair of vectors \( v \) and \( v' \), the numbers

\[
\langle Sv, v' \rangle \quad \text{and} \quad \langle v, Tv' \rangle,
\]

respectively. (107)

Our task is to determine their matrices in a given basis \( \mathcal{B} \). Below \( Q \) is the matrix of \( \langle \cdot, \cdot \rangle \) in basis \( \mathcal{B} \), cf. (31). To simplify notation, \( A_S, A_T \) and \( \begin{bmatrix} \cdot \end{bmatrix}_\mathcal{B} \) stand for: \( A_S(\mathcal{B}) \), \( A_T(\mathcal{B}) \) and \( \begin{bmatrix} \cdot \end{bmatrix}_\mathcal{B} \), respectively.

According to (36), one has

\[
\langle Sv, v' \rangle = \begin{bmatrix} v \end{bmatrix}^\top Q \begin{bmatrix} v' \end{bmatrix} = (A_S [v])^\top Q [v'] = \begin{bmatrix} v \end{bmatrix}^\top (A_S^\top Q) [v']
\]

and

\[
\langle v, Tv' \rangle = \begin{bmatrix} v \end{bmatrix}^\top Q T [v] = \begin{bmatrix} v \end{bmatrix}^\top Q (A_T [v']) = \begin{bmatrix} v \end{bmatrix}^\top (Q A_T) [v'].
\]
4.3.2 The matrix of the adjoint operator

We say that $S$ and $T$ are adjoint to each other with respect to the bilinear pairing $\langle , \rangle$ if two pairings introduced above coincide:

$$\langle Sv, v' \rangle = \langle v, Tv' \rangle \quad (v, v' \in V).$$

Our calculations express this in terms of the matrices of the corresponding bilinear pairings in basis $B$:

$$A^T S Q = Q A_T \quad (108)$$

or, equivalently,

$$Q^T A_S = A_T^T Q^T. \quad (109)$$

If $\langle , \rangle$ is nondegenerate, then $Q$ is invertible and

$$A_S = (Q^T)^{-1} A_T^T Q^T. \quad (110)$$

4.3.3

If $\langle , \rangle$ is a scalar product, then $Q$ is symmetric and (110) becomes

$$A_S = Q^{-1} A_T^T Q \quad (111)$$

or, using the adjoint operator notation $T^*$,

$$A_{T^*} = Q^{-1} A_T^T Q. \quad (112)$$

4.3.4

The basis is orthonormal if and only if $Q$ is the identity matrix. In this case, the matrix of the adjoint operator is simply the transpose of the matrix of $T$:

$$A_{T^*} = (A_T)^\tau. \quad (113)$$

4.3.5

Note that (113) agrees with the fact that

$$(A^T x) \cdot y = x \cdot (Ay).$$
4.3.6 Selfadjoint operators

An operator on a vector space equipped with a scalar product $\langle \, , \rangle$ is said to be selfadjoint if $T = T^*$. Equivalently,

$$\langle Tv, v' \rangle = \langle v, Tv' \rangle \quad (v, v' \in V). \quad (114)$$

4.3.7

The operator is selfadjoint if and only if its matrix $A_T$ in any orthonormal basis is symmetric.

**Exercise 11** Show that the orthogonal projection $P$ onto a subspace $W$ is selfadjoint.

**Exercise 12** Let $W$ be an invariant subspace of a selfadjoint operator $T$. Show that also its orthogonal complement $W^\perp$ is an invariant subspace.

**Exercise 13** Let $v$ and $v'$ be eigenvectors of a selfadjoint operator $T$ with eigenvalues $\lambda \neq \lambda'$. Show that $v \perp v'$.

4.3.8

Any selfadjoint operator on a finite-dimensional nonzero vector space has at least one eigenvalue.

4.3.9

Let us accept this fact and see what does follow from it. The vector subspace $V'$ spanned by all the eigenvectors, cf. (94), is obviously invariant. Since also $(V')^\perp$ is invariant and $T$ restricted to $(V')^\perp$ is a selfadjoint operator on $(V')^\perp$, it has a nonzero eigenvector in $(V')^\perp$ as long as $(V')^\perp$ is nonzero. But all such eigenvectors belong to $V'$. Thus, $(V')^\perp$ is zero and $V' = V$.

4.3.10

We deduced that $V$ is the direct sum of orthogonal eigenspaces of the selfadjoint operator $T$. By choosing an orthogonal basis in each eigenspace, we arrive at the following fundamentally important result.
Theorem 4.1 (Spectral Theorem) Any selfadjoint operator $T$ is diagonalizable with a basis consisting of orthonormal eigenvectors. More precisely, $T$ can be represented as

$$T = \sum_{1 \leq i \leq n} \lambda_i u_i u_i^* \quad (115)$$

where $\{u_1, \ldots, u_n\}$ is an orthonormal basis of $V$ consisting of eigenvectors $u_i$ with the eigenvalues $\lambda_i$. Here each eigenvalue is encountered as many times as the dimension of the eigenspace $V_{\lambda_i}$.

4.3.11

Any diagonalizable operator is selfadjoint with respect to the dot product $\cdot_B$ where $B$ denotes any basis for which the matrix $A_T(B)$ is diagonal. It follows that an operator is diagonalizable if and only if it is selfadjoint with respect to at least one scalar product.