MAS4107 Linear Algebra 2

Peter Sin
University of Florida
email: sin@math.ufl.edu

General Prerequisites

Familiarity with the notion of mathematical proof and some experience in reading and writing proofs. Familiarity with standard mathematical notation such as summations and notations of set theory.

Linear Algebra Prerequisites

Familiarity with the notion of linear independence. Gaussian elimination (reduction by row operations) to solve systems of equations. This is the most important algorithm and it will be assumed and used freely in the classes, for example to find coordinate vectors with respect to basis and to compute the matrix of a linear map, to test for linear dependence, etc. The determinant of a square matrix by cofactors and also by row operations.

0. Introduction

These notes include some topics from MAS4105, which you should have seen in one form or another, but probably presented in a totally different way. They have been written in a terse style, so you should read very slowly and with patience. Please feel free to email me with any questions or comments. The notes are in electronic form so sections can be changed very easily to incorporate improvements.

1. Groups and Fields

1.1. Groups

The theory of groups is an important and interesting area of mathematics in its own right, but we introduce the concept here as an aid to understanding the concepts of fields and later vector spaces, the main subjects of this course.

Definition 1.1. A group is a set G with a binary operation (which we indicate by $*$ here) satisfying the following axioms:

1. (Identity element) There exists an element $e \in G$ such that for all $g \in G$ we have $e * g = g * e = g$.
2. (Associativity) For any three elements $f, g, h \in G$ we have $(f * g) * h = f * (g * h)$.
3. (Inverses) For each element $g \in G$, there exists an element $g' \in G$ such that $g * g' = e = g' * g$.

Exercise 1.1. Show that a group has a unique identity element. Show that for each element g in a group G there is exactly one element which satisfies the properties of g' in the Inverses axiom.

Definition 1.2. A binary operation is called *commutative* if the additional property holds that for any two elements g and h, we have $g * h = h * g$. A group whose operation is commutative is often called an *abelian* group.

Most of the groups we will consider are abelian, including the following examples.

Example 1.2. Let $G = \mathbb{Z}$ and let $*$ be addition Check that this is a group. Which integer is e? Given an integer, what is it's inverse element in this group?

Example 1.3. Let G be the set of nonzero positive real numbers and let $*$ be multiplication. Is this a group?

Example 1.4. Let G be the set of positive integers. Is this a group under either addition or multiplication?

1.2. Fields

Fields are algebraic systems with many properties similar to the set $\mathbb{R}$ of real numbers, which is an example of a field. In linear algebra, fields play the role of 'scalars'. Most of the basic theory for vectors and matrices with real entries holds over arbitrary fields, and it is often profitable to work in the more general context of fields.

The definition of fields involves two binary operations, which are usually called addition and multiplication and indicated with notation $\alpha + \beta$ for the sum and $\alpha\beta$ for the product of two elements. You already know some examples, such as the field $\mathbb{R}$ of real numbers and the field $\mathbb{Q}$ of rational numbers, and it is an easy exercise to check that these satisfy the following definition.

Definition 1.3. A field is a set F which has two commutative binary operations, called addition and multiplication such that:

1. F is a group under addition.
2. $F \setminus \{0\}$ is a group under multiplication, where 0 is the identity element of the additive group.
3. (Distributive Law) For all elements α, β, $\gamma \in F$ we have $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Remark 1.1. The identity element of the multiplicative group is usually denoted by 1. The Distributive Law is the axiom which ties together the two binary operations.

The existence of inverses for addition means that we have subtraction and the existence of inverses for multiplication means we have division (by elements other than 0). So a field is basically an algebraic system where one can perform all four of the usual operations of arithmetic and the familiar properties such as the associative, commutative and distibutive laws hold.

Exercise 1.5. Determine which of the following are fields (using the usual addition and multiplication): $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{C}$.

Example 1.6. Consider the set with two elements $\overline{0}$ and $\overline{1}$. The recipe for addition and multiplication is to think of $\overline{0}$ as "even" and $\overline{1}$ as "odd". Then, since the sum of two odd integers is even, we have $\overline{1} + \overline{1} = \overline{0}$, and similarly $\overline{1}.\overline{1} = \overline{1}$, etc. Check that this is a field. For the more ambitious, show that apart from renaming the elements, the above addition and multiplication are the only way to have a field with two elements.

2. Vector Spaces

2.1. Vectors

What is a vector? In courses on analytic geometry, vectors in the plane or in space are often described as arrows and represent physical quantities having magnitude and direction. This is certainly an good intuitive way to think of vectors in euclidean space, and it is how vectors first arose and how they are often applied. However, many students find it difficult to connect this idea of vectors with the more general algebraic definition given in linear algebra courses. If this is the case it may be better to temporarily drop the imprecise geometric intuition until you are comfortable working with the algebraic axioms, and remember that a vector is simply an element in a special kind of abelian group called a vector space, no more, no less. So, once we have the definition of vector spaces we will know what vectors are. The definition of vector spaces involves two sets, an abelian group V and a field F. The elements of V are called *vectors*, and those of F are called *scalars*. The group operation in V is written as addition. We also have addition and multiplication in F. (Note that the "+" sign is used for both additions, although they are not related.)In a vector space, there is also a notion of *scalar multiplication* of vectors, namely, a way of combining each $v \in V$ and $\alpha \in F$ to give a new vector denoted αv.

Definition 2.1. A vector space over a field F is an abelian group V, equipped with a scalar multiplication such that the following properties hold:

1. $\alpha(v + w) = \alpha v + \alpha w$, $\forall v, w \in V$, $\forall \alpha \in F$.
2. $\alpha(\beta v) = (\alpha\beta)v$, $\forall v \in V$, $\forall \alpha, \beta \in F$.
3. $(\alpha + \beta)v = \alpha v + \beta v$, $\forall v \in V$, $\forall \alpha, \beta \in F$.
4. $1v = v$, where 1 is the multiplicative identity of F.

Notation 2.2. The additive identity in V will be denoted by $\mathbf{0}$, using bold type to distinguish this vector from the scalar element 0.

Exercise 2.1. Prove that for all $v \in V$ we have $0v = \mathbf{0}$.

Remark 2.1. Our convention is that field elements (scalars) multiply vectors from the left. So, the symbol $v\alpha$ is meaningless at this point.

2.2. Examples

You should check your understanding of all new concepts against the list of assorted examples in this subsection.

Example 2.2. F^n, n-tuples of elements of F, with entrywise addition and scalar multiplication.

Example 2.3. $\mathrm{Mat}_{m \times n}(F)$, matrices with m rows and n columns with entries from the field F. Addition and scalar muliplication are entrywise.

Example 2.4. The space $\mathrm{Poly}(F)$ of polynomials over F. This is the set of all expressions of the form

$$p(t) = \alpha_0 + \alpha_1 t + \cdots \alpha_d t^d,$$

where $d \in \mathbb{N}$ and $\alpha_i \in F$. Two polynomials are added by adding the coefficients of like powers of t. Scalar multiplication simply multiplies every term by the given scalar. It is important to realize that we are not thinking of polyniomials as functions here. The variable t here is just a placeholder. Therefore two polynomials are equal iff they have the same coefficients. For example if F is the field with two elements, the *functions* t^2 and t are the same. To see this just plug in $\overline{0}$ and $\overline{1}$. But t and t^2 are considered distinct polynomials.

Example 2.5. $\mathrm{Poly}_k(F)$, the subset of $\mathrm{Poly}(F)$ consisting of polynomials of degree at most k, with the same rules for addition and scalar multiplication as for $\mathrm{Poly}(F)$.

Example 2.6. Let X be a set and F^X the set of all functions from X to F. The sum of two functions is defined as $(f+g)(x) = f(x)+g(x)$ and scalar multiplication by $(\alpha f)(x) = \alpha f(x)$.

Exercise 2.7. Discuss which of the examples can be regarded as special cases of example 2.6, by identifying X in each case.

Notation 2.3. For economy of language, we adopt the convention that unless otherwise stated, vector spaces will be over the field F and denoted by Roman capitals V, W, etc. Vectors will be denoted by lower case roman letters v, v', w, etc. and sclars by lower case Greek letters α, β, etc.

3. Subspaces, Linear Dependence and Generation

3.1. Linear combinations

In a vector space V addition of vectors and scalar multiplication of vectors both result in vectors. Starting with vectors $v_1, \ldots, v_k$ a vector of the form

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k, \quad \text{where } \alpha_i \in F, \tag{1}$$

is called a *linear combination* of the vectors $v_1, \ldots, v_k$. The scalars α_i may be any elements of F, including 0. More generally, if S is any set of vectors, the linear combinations of S are defined to be the linear combinations of finite subsets of S.

Exercise 3.1. Let S be a set of vectors (not necessarily finite). Show that any linear combination of linear combinations of S is a linear combination of S.

Definition 3.1. Let S be a subset of a vector space V. The *linear span* of S, denoted $\langle S \rangle$, is the set of all linear combinations of S. (By convention, we take $\langle \emptyset \rangle$ to be $\{\mathbf{0}\}$.)

3.2. Subspaces

Definition 3.2. A subset W of a vector space V over F is a *subspace* if it is a vector space over F under the same addition and scalar multiplication.

Exercise 3.2. A non-empty W subset of V is a subspace if and only if it it contains the sum of any two elements of W (closed under addition) and all scalar multiples of elements of W (closed under scalar multiplication).

Exercise 3.3. Let S be any subset of V. Then $\langle S \rangle$ is a subspace and any subspace which contains S contains $\langle S \rangle$.

3.3. Linear dependence

Let $v_1, \ldots, v_k \in V$ ($k \geq 1$). Consider the problem of expressing $\mathbf{0}$ as a linear combination (1) of $v_1, \ldots, v_k \in V$. An obvious solution would be to choose all scalars α_i to be 0, so the question is meaningful only if this trivial case is excluded.

Definition 3.3. The vectors $v_1, \ldots, v_k \in V$ are *linearly dependent* if there exist scalars $\alpha_1, \ldots, \alpha_k$ such that

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \mathbf{0}, \tag{2}$$

with at least one of the α_i not equal to 0.

More generally, a set S of vectors is linearly dependent if there is a finite subset of S which is linearly dependent.

Definition 3.4. A set S of vectors in V is *linearly independent* if it is not linearly dependent.

4. Bases and Coordinates

4.1. Bases

Definition 4.1. A subset $\mathcal{B}$ of V is a *basis* of V iff

1. $\mathcal{B}$ is linearly independent; and
2. The linear span of $\mathcal{B}$ is equal to V.

Lemma 4.1. *Suppose $\mathcal{B}$ is a basis of V. Then each $v \in V$ may be expressed as a linear combination of $\mathcal{B}$. Furthermore, this expression is unique (up to addition or deletion of terms consisting of vectors in $\mathcal{B}$ multiplied by 0).*

Proof. Exercise. □

Definition 4.2. A subset S of V with the property that $\langle S \rangle = V$ is called a *generating set* for V.

Definition 4.3. V is *finite-dimensional* if it has a finite generating set. If not, we say V is *infinite-dimensional*

Lemma 4.2. *Every finite-dimensional vector space has a basis.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a generating set of smallest possible size. If it were not linearly independent, then one of the v_i would be a linear combination of the others and the set obtained by removing it would still be a generating set, contradicting the minimality. □

Remark 4.1. The existence of bases in an arbitrary vector space depends on the Axiom of Choice.

Lemma 4.3. (Exchange Lemma)*Suppose* $v_1, \ldots, v_n$ *form a generating set for* V *and* $x_1, \ldots, x_m$ *are linearly independent. Then* $n \geq m$ *and there are* m *of the* v_i *(call them* $v_1, \ldots, v_m$*) which can be replaced by the* x_i *so that the resulting set* $x_1, \ldots, x_m, v_{m+1}, \ldots, v_n$ *form a generating set.*

Proof. By induction on m, the case $m = 0$ being trivial. Assume true for $m - 1$. Then since $x_1, \ldots, x_{m-1}$ are linearly independent, the induction hypothesis tells us that $m - 1 \leq n$ and allows us to number the v_i in a way that $x_1, \ldots, x_{m-1}, v_m, \ldots, v_n$ form a generating set. Therefore x_m can be written as a linear combination

$$x_m = \sum_{i=1}^{m-1} \alpha_i x_i + \sum_{i=m}^{n} \beta_i v_i \tag{3}$$

Now since $x_1, \ldots, x_m$ are linearly independent, at least one β_i must be nonzero, so $n \geq m$. Renumbering the v_i if necessary, we can assume $\beta_m \neq 0$. The lemma will be established if we show that $x_1, \ldots, x_m, v_{m+1}, \ldots, v_n$ is a generating set. Since we know by the inductive hypothesis that $x_1, \ldots, x_{m-1}, v_m, \ldots, v_n$ form a generating set, it is enough to show that v_m is a linear combination of $x_1, \ldots, x_m, v_{m+1}, \ldots, v_n$. But this follows by rearranging equation (3) since we can divide by the nonzero coefficient β_m. □

Corollary 4.4. *In a finite-dimensional vector space all bases have the same number of elements.*

Proof. Let $\mathcal{B}$ and $\mathcal{B}'$ be bases of size n and m respectively. Then since $\mathcal{B}$ is a generating set and $\mathcal{B}'$ is linearly independent, we have $m \leq n$ be the Exchange Lemma. Now, interchanging the roles of $\mathcal{B}$ and $\mathcal{B}'$ shows $n \leq m$ also. □

Exercise 4.1. Prove that in a finite-dimensional vector space any linearly independent set can be expanded to a basis and any generating set can be contracted to a basis.

Definition 4.4. *Dimension.* If V is a finite-dimensional vector space over F then its *dimension*, denoted $\dim V$ is defined to be the common size of all its bases.

4.2. Coordinates

Suppose $\dim V = n$ and let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis. Then by Lemma 4.1, each vector v determines a unique n-tuple of scalars by the formula

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Definition 4.5. The tuple $(\alpha_1, \dots, \alpha_n)$ determined by v is denoted $[\, v\,]_{\mathcal{B}}$ and called the *coordinate vector* of v with respect to the ordered bases $\mathcal{B}$. Coordinate vectors are to be thought of as column vectors, but sometimes written as row vectors to save space. This abuse of notation is safe as long as one is careful in places where more precision is needed.

Conversely, each n-tuple of scalars determines a unique vector, also by (4.2). In other word, the choice of an ordered bases $\mathcal{B}$ produces a one-one correspondence $\phi_{\mathcal{B}} : V \to F^n$, sending v to $[\, v\,]_{\mathcal{B}}$.

Definition 4.6. The map $\phi_{\mathcal{B}}$ is called is called the *coordinate map* with respect to $\mathcal{B}$.

Note that if we start with an element of F^n, we can find the corresponding vector in v simply by substituting in (4.2). However, if we start with a vector, we have to solve a system of linear equations in order to compute its coordinate vector.

Notation 4.7. Let e_i the element of F^n with 1 in the i-th entry and 0 in all other entries.

Exercise 4.2. Show that for the i-th elelemnt v_i in an ordered basis $\mathcal{B}$ we have $[\, v_i\,]_{\mathcal{B}} = e_i$.

5. Linear Maps and Matrices

5.1. Linear maps

Definition 5.1. Let V and W be vector spaces over F. A map $T : V \to W$ is *linear* iff it satisfies the conditions:

1. $T(v + v') = Tv + Tv'$, $\forall v, v' \in V$.
2. $T(\alpha v) = \alpha Tv$, $\forall v \in V$, $\forall \alpha \in F$.

Example 5.1. The identity map $\mathrm{id} : V \to V$, defined by $\mathrm{id}(v) = v$, $\forall v \in V$ is a trivial example, but one which will be important in the next Chapter.

Example 5.2. If $T : V \to W$ and $S : W \to Y$ are linear maps, then the composite $S \circ T : V \to Y$ defined by $(ST)(v) = S(T(v))$ is a linear map.

Example 5.3. The coordinate map $\phi_{\mathcal{B}} : V \to F^n$ with respect to a basis $\mathcal{B}$ is a a linear map.

Exercise 5.4. Prove that if a linear map $T : V \to W$ is a one-one correspondence, then the inverse mapping $T^{-1} : W \to V$ is also linear.

Definition 5.2. A bijective linear map is called an *isomorphism*. Vector spaces related by an isomorphism are said to be *isomorphic*.

Exercise 5.5. Prove that the coordinate map $\phi_{\mathcal{B}} : V \to F^n$ is an isomorphism. Prove that two finite-dimensional vector spaces over F are isomorphic if and only of they have the same dimension.

Theorem 5.1. *Let V and W be vector spaces over F. Suppose $v_1, \ldots, v_n$ is a basis of V. Then, given any n vectors $w_1, \ldots, w_n$ in W, there exists a unique linear map $T : V \to W$ such that $Tv_i = w_i$, for $i = 1, \ldots, n$.*

Proof. The uniquenes of T, if it exists, follows from the fact that two linear maps which agree on a set of vectors must agree on all linear cobinations of that set. It remains to show that T exists. Define $S : F^n \to W$ by $S(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^n \alpha_i w_i$. This map is easily checked to be linear. Since we are given a basis $\mathcal{B}$ of V, we have a coordinate map $\phi_\mathcal{B} : V \to F^n$. Let T be the composite map $S \circ \phi_\mathcal{B} : V \to W$. By Exercise 4.2) and the definiton of S we have so $Tv_i = S(\phi_\mathcal{B}(v_i)) = S(e_i) = w_i$ as required. □

5.2. Matrices

Definition 5.3. Let V and W be vector spaces over F Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V and $\mathcal{C} = \{w_1, \ldots, w_m\}$ a basis of W. Let $T : V \to W$ be a linear map. We define the matrix $[T]_\mathcal{B}^\mathcal{C}$ of T with respect to these bases (and their given numbering) to be the $m \times n$ matrix whose j-th column is equal to $[Tv_j]_\mathcal{C}$, for $j = 1, \ldots, n$.

Remark 5.1. If we renumber the elements of $\mathcal{B}$, the matrix we will obtain will not be quite the same, the columns will be permuted. Likewise, renumbering $\mathcal{C}$ results in a row permutation. So to specify the matrix exactly, the ordering of the basis must be specified.

Exercise 5.6. Suppose $\dim V = n$ and $\dim W = m$ and let ordered bases $\mathcal{B}$ and $\mathcal{C}$ be given. Show that every $m \times n$ matrix with entries in F occurs as the matrix $[T]_\mathcal{B}^\mathcal{C}$ for some linear map $T : V \to W$. (Hint: Theorem 5.1)

Definition 5.4. *Matrix multiplication.* Let A be an $m \times n$ matrix with entries in F and let B be an $n \times p$ matrix. The product AB is defined to be the $m \times p$ matrix whose (i, j) entry is $\sum_{k=1}^n a_{ik} b_{kj}$, for $i = 1, \ldots, m$ and $j = 1, \ldots, p$. (The matrix product is not defined unless the number of columns of A equals the number of rows of B.)

Example 5.7. A column vector of length n, for instance a coordinate vector, may be viewed as an $n \times 1$ matrix and hence multiplied on the left by an $m \times n$ matrix.

Exercise 5.8. Show that multiplying column vectors on the left by a fixed $m \times n$ matrix A is a linear map from F^n to F^m.

Exercise 5.9. Think of the columns of B as column vectors. Show that the j-th column of AB is the matrix product of A with the j-th column of B.

Theorem 5.2. *Let $T : V \to W$ be a linear map and let $\mathcal{B}$ be an ordered basis of V and $\mathcal{C}$ one of W. Then for all $v \in V$ we have*

$$[\,Tv\,]_{\mathcal{C}} = [\,T\,]^{\mathcal{C}}_{\mathcal{B}}[\,v\,]_{\mathcal{B}} \tag{4}$$

Proof. Consider first the composite map $\phi_{\mathcal{C}} \circ T : V \to F^m$, where $m = \dim W$. The left hand side of the equation is the image of v under this map. The right hand side is the image of v under the composite map consisting of $\phi_{\mathcal{B}} : V \to F^n$ followed by the map $F^n \to F^m$ given by left multiplication by the matrix $[\,T\,]^{\mathcal{C}}_{\mathcal{B}}$. Since both these composite maps are linear maps from V to F^m, they will be equal if they agree on the elements of the basis $\mathcal{B}$. Let v_i be the i-th element of $\mathcal{B}$. Then $\phi_{\mathcal{C}}(T(v_i))$ is the i-th column of $[\,T\,]^{\mathcal{C}}_{\mathcal{B}}$, by Definition 5.3. On the other hand

$$[\,T\,]^{\mathcal{C}}_{\mathcal{B}}[\,v_i\,]_{\mathcal{B}} = [\,T\,]^{\mathcal{C}}_{\mathcal{B}} e_i,$$

which is also equal to the i-th column of $[\,T\,]^{\mathcal{C}}_{\mathcal{B}}$. □

It is helpful to think of Theorem 5.2 as saying that in the following diagram both ways to go from top left to bottom right give the same answer. The diagram is said to be commutative.

$$\begin{tikzcd} \mathcal{B} & \mathcal{C} \\ V \arrow[r, "T"] \arrow[d, "\phi_{\mathcal{B}}"] & W \arrow[d, "\phi_{\mathcal{C}}"] \\ F^n \arrow[r, "{[\,T\,]_{\mathcal{B}}^{\mathcal{C}}}"] & F^m \end{tikzcd} \tag{5}$$

5.3. Composition of maps and matrices

Theorem 5.3. *Let $T : V \to W$ and $S : W \to Y$ be linear maps and let $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ be (ordered) bases of V, W and Y respectively. Then we have*

$$[\,S \circ T\,]_{\mathcal{B}}^{\mathcal{D}} = [\,S\,]_{\mathcal{C}}^{\mathcal{D}} [\,T\,]_{\mathcal{B}}^{\mathcal{C}}. \tag{6}$$

Proof. By definition, the i-th column of $[\,S \circ T\,]_{\mathcal{B}}^{\mathcal{D}}$ is $[\,S(T(v_i))\,]_{\mathcal{D}}$, where v_i is the i-th element of $\mathcal{B}$. By Theorem 5.2, and the definition of matrix multiplcation, we have

$$[\,S(T(v_i))\,]_{\mathcal{D}} = [\,S\,]_{\mathcal{C}}^{\mathcal{D}} [\,T(v_i)\,]_{\mathcal{C}} = [\,S\,]_{\mathcal{C}}^{\mathcal{D}} (i\text{-th column of } [\,T\,]_{\mathcal{B}}^{\mathcal{C}}) = i\text{-th column of } [\,S\,]_{\mathcal{C}}^{\mathcal{D}} [\,T\,]_{\mathcal{B}}^{\mathcal{C}}.$$

$\square$

Let us interpret this in terms of commutative diagrams. Consider the diagram obtained by combining two versions of (5).

Introduction
Groups and Fields
Vector Spaces
Subspaces, Linear . . .
Bases and Coordinates
Linear Maps and . . .
Change of Basis
More on Linear Maps
Linear Endomorphisms
Quotient Spaces
Spaces of Linear . . .
Direct Sums
Minimal polynomial
Bilinear Forms
Hermitian Forms
Euclidean and . . .
Self-Adjoint Linear . . .
Notation
Back
Full Screen
Close
Quit

$$
\begin{tikzcd}
\mathcal{B} & \mathcal{C} & \mathcal{D} \\
V \arrow[r, "T"] \arrow[d, "\phi_{\mathcal{B}}"] & W \arrow[r, "S"] \arrow[d, "\phi_{\mathcal{C}}"] & Y \arrow[d, "\phi_{\mathcal{D}}"] \\
F^n \arrow[r, "{[T]_{\mathcal{B}}^{\mathcal{C}}}"] & F^m \arrow[r, "{[S]_{\mathcal{C}}^{\mathcal{D}}}"] & F^p
\end{tikzcd}
\tag{7}
$$

Theorem 5.3 says that the matrix of the composite of the two maps in the top row is obtained by mutliplying the two matrices in the bottom row together. This can be extended to composites of three or more maps.

Exercise 5.10. (Harder) Use the above observation (applied to three maps) together with Exercise 5.6 and the fact that composition of maps is associative to prove that matrix multiplication is associative, i.e., if A, B and C are matrices whose shapes allow them to be multiplied, then $(AB)C = A(BC)$.

Introduction
Groups and Fields
Vector Spaces
Subspaces, Linear ...
Bases and Coordinates
Linear Maps and ...
Change of Basis
More on Linear Maps
Linear Endomorphisms
Quotient Spaces
Spaces of Linear ...
Direct Sums
Minimal polynomial
Bilinear Forms
Hermitian Forms
Euclidean and ...
Self-Adjoint Linear ...
Notation

Back
Full Screen
Close
Quit

6. Change of Basis

If we have a linear map $T : V \to W$ and bases $\mathcal{B}$ and $\mathcal{C}$ of V and W respectvely, then we have seen how to compute the coordinates of vectors and the matrix of the map. Now we want to consider what happens if we choose different bases. Of course, we can compute coordinates and matrices just as easily in the new basis, but the point is to understand how the coordinate vectors and matrix in the new bases are related to the corresponding objects computed in the old bases. A proper understanding of this material is the key to applications of coordinates in many fields including geometry and engineering. We already have the two general formulae Theorem 5.2 and Theorem 5.3 needed for this pupose. But these formulae are so general that some skill is needed in to apply them effectively. In particular, one must make the right choice of V, W, $\mathcal{B}$, $\mathcal{C}$, T, etc. to suit each particular proof. The diagrams of the previous chapter provide a convenient notation indicating these choices.

6.1. Change of Coordinates

We begin by looking at the effect of changing bases on the coordinates of a vector. Let $\mathcal{B}$ and $\mathcal{B}'$ be two bases. We want to compare $[\,v\,]_{\mathcal{B}}$ with $[\,v\,]_{\mathcal{B}'}$ for all $v \in V$. To do this, we (5.2) with $W = V$ $T = \mathrm{id}$, and $\mathcal{C} = \mathcal{B}'$, which gives

Theorem 6.1. (Change of coordinates formula)

$$[\,v\,]_{\mathcal{B}'} = [\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}[\,v\,]_{\mathcal{B}}. \tag{8}$$

We wish to study the matrix $[\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}$ and its counterpart $[\,\mathrm{id}\,]_{\mathcal{B}'}^{\mathcal{B}}$ further.

Back
Full Screen
Close
Quit

Equation (8) corresponds to the diagram

$$
\begin{tikzcd}
\mathcal{B} & \mathcal{B}' \\
V \arrow[r, "\mathrm{id}_V"] \arrow[d, "\phi_{\mathcal{B}}"] & V \arrow[d, "\phi_{\mathcal{B}'}"] \\
F^n \arrow[r, "{[\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}}"] & F^n
\end{tikzcd}
\tag{9}
$$

and there is also a similar diagram with $\mathcal{B}$ and $\mathcal{B}'$ swapped. Combining these, we have

$$
\begin{tikzcd}
\mathcal{B}' & \mathcal{B} & \mathcal{B}' \\
V \arrow[r, "\mathrm{id}_V"] \arrow[d, "\phi_{\mathcal{B}'}"] & V \arrow[r, "\mathrm{id}_V"] \arrow[d, "\phi_{\mathcal{B}}"] & V \arrow[d, "\phi_{\mathcal{B}'}"] \\
F^n \arrow[r, "{[\,\mathrm{id}\,]_{\mathcal{B}'}^{\mathcal{B}}}"] & F^n \arrow[r, "{[\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}}"] & F^n
\end{tikzcd}
\tag{10}
$$

The compositon of the top row is the identity, so by Theorem 5.3 we have

$$
[\,\mathrm{id}_V\,]_{\mathcal{B}}^{\mathcal{B}'}[\,\mathrm{id}_V\,]_{\mathcal{B}'}^{\mathcal{B}} = [\,\mathrm{id}_V\,]_{\mathcal{B}'}^{\mathcal{B}'} = \mathrm{I}_n. \tag{11}
$$

By symmetry we also have $[\,\mathrm{id}_V\,]_{\mathcal{B}'}^{\mathcal{B}}[\,\mathrm{id}_V\,]_{\mathcal{B}}^{\mathcal{B}'} = \mathrm{I}_n$.

Definition 6.1. Two $n \times n$ matrices A and B are said to be inverse to each other if $AB = \mathrm{I}_n = BA$. It is left as an exercise to show that the inverse of a matrix is unique if it exists. If A has an inverse, we write it as A^{-1}.

Back
Full Screen
Close
Quit

We have proved the following result.

Lemma 6.2. *With the same notation as above, the matrices* $[\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}$ *and* $[\,\mathrm{id}\,]_{\mathcal{B}'}^{\mathcal{B}}$ *are inverse to each other.*

6.2. Change of bases and matrices

Let $\mathcal{B}$ and $\mathcal{B}'$ be two bases of the n-dimensional vector space V and let $\mathcal{C}$ and $\mathcal{C}'$ be two bases of the m-dimensional vector space W. Let $T : V \to W$ be a linear map. Our goal is to relate $[\,T\,]_{\mathcal{B}'}^{\mathcal{C}'}$ and $[\,T\,]_{\mathcal{B}}^{\mathcal{C}}$

Theorem 6.3. *With the above notation, we have*

$$[\,T\,]_{\mathcal{B}'}^{\mathcal{C}'} = [\,\mathrm{id}_W\,]_{\mathcal{C}}^{\mathcal{C}'}[\,T\,]_{\mathcal{B}}^{\mathcal{C}}[\,\mathrm{id}_V\,]_{\mathcal{B}'}^{\mathcal{B}}. \tag{12}$$

Proof. This follows by applying Theorem 5.3 to the following diagram. It is a good test of understanding of this material to try to fill in the detailed reasoning.

$$\begin{tikzcd}
\mathcal{B}' & \mathcal{B} & \mathcal{C} & \mathcal{C}' \\
V \arrow[r, "\mathrm{id}_V"] \arrow[d, "\phi_{\mathcal{B}'}"] & V \arrow[r, "T"] \arrow[d, "\phi_{\mathcal{B}}"] & W \arrow[r, "\mathrm{id}_W"] \arrow[d, "\phi_{\mathcal{C}}"] & W \arrow[d, "\phi_{\mathcal{C}'}"] \\
F^n \arrow[r, "{[\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}}"] & F^n \arrow[r, "{[\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{C}}}"] & F^m \arrow[r, "{[\,\mathrm{id}\,]_{\mathcal{C}}^{\mathcal{C}'}}"] & F^m
\end{tikzcd} \tag{13}$$

□

Back
Full Screen
Close
Quit

Theorem 6.3 is very general. it holds for all choices of vector spaces, bases and linear maps. Many applications involve special cases where some of these choices are the same. As an example we give the important case $V = W$, $\mathcal{B} = \mathcal{C}$, $\mathcal{B}' = \mathcal{C}'$. Let $A = [\,\mathrm{id}\,]_{\mathcal{B}}^{\mathcal{B}'}$. Then $[\,\mathrm{id}\,]_{\mathcal{B}'}^{\mathcal{B}} = A^{-1}$ by Lemma 6.2, so we obtain the following result.

Corollary 6.4. *Let $T : V \to V$ be a linear map and let $\mathcal{B}$ and $\mathcal{B}'$ be two bases of V. Then*

$$[\,T\,]_{\mathcal{B}'}^{\mathcal{B}'} = A[\,T\,]_{\mathcal{B}}^{\mathcal{B}}A^{-1}. \tag{14}$$

Introduction
Groups and Fields
Vector Spaces
Subspaces, Linear . . .
Bases and Coordinates
Linear Maps and . . .
Change of Basis
More on Linear Maps
Linear Endomorphisms
Quotient Spaces
Spaces of Linear . . .
Direct Sums
Minimal polynomial
Bilinear Forms
Hermitian Forms
Euclidean and . . .
Self-Adjoint Linear . . .
Notation
Back
Full Screen
Close
Quit

7. More on Linear Maps

7.1. The kernel and image of a linear map

Definition 7.1. Let $T : V \to W$ be a linear map. The *kernel* of T is the set

$$\text{Ker } T = \{v \in V \mid Tv = \mathbf{0}_W\}. \tag{15}$$

The *image* of T is the set

$$\text{Im } T = \{w \in W | \exists v \in V, w = Tv\} \tag{16}$$

Exercise 7.1. Ker T is a subspace of V and Im T is a subspace of W.

Exercise 7.2. Let $T : V \to W$ be a linear map, Suppose V and W have dimensions n and m respectively. Let $v_1, \dots v_k$ form a basis for Ker T.
(a) Explain why we can expand this to a basis $v_1, \dots, v_k, \dots, v_n$ of V.
Set $w_1 = Tv_{k+1}, \dots, w_{n-k} = Tv_n$.
(b) Prove that $w_1, \dots, w_{n-k}$ form a basis of Im T.
(c) Deduce that $\dim \text{Ker } T + \dim \text{Im } T = \dim V$.
(d) Let $\mathcal{B}$ be the above basis of V and let $\mathcal{C}$ be a basis of W obtained by expanding the above basis of Im T. Compute the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$.
(e) Deduce that given any $m \times n$ matrix, there exist invertible matrices P $(m \times m)$ and Q $(n \times n)$ such that PAQ has the simple form of the matrix in (d).

Definition 7.2. The dimension of Im T is called the *rank* of T and the dimension of Ker T is called the *nullity* of T.

8. Linear Endomorphisms

In this section we shall focus on the special situation of a linear map $T : V \to V$ from a vector space V to itself. Such linear maps are called *endomorphisms* and the set of all endomorphisms of V is denoted by $\mathrm{End}(V)$. Here, and in other texts, there will be references to "the matrix of T with respect to the basis $\mathcal{B}$". Of course the general definiton of the matrix of T requires us to specify two bases, not just one. But in the context of endomorphisms, it means that one should take the same basis twice, for V in its roles as the domain and as the codomain of T.

We have already seen that the matrix of T in one basis will be similar to the matrix of T in any other basis. One of the objectives of this theory is to pick out of all these similar matrices, special ones which have a particularly simple form. As we have seen, this is equivalent to finding bases which are especially compatible with T.

8.1. Invariant subspaces

Definition 8.1. A subspace U of V is said to be T-*invariant* iff $T(U) \subseteq U$. For example, V, $\{\mathbf{0}\}$, $\mathrm{Ker}\, T$ and $\mathrm{Im}\, T$ are all T-invariant. (Exercise: Prove the last two.)

Exercise 8.1. Suppose U is a T-invariant subspace of V. Let $\mathcal{B}$ be a basis of V obtained by expanding a basis of U. Show that the matrix of T in this basis has the form

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

where A is an $r \times r$ submatrix ($r = \dim U$), 0 is an $(n-r) \times r$ submatrix of zeros and B and C are submatrices of sizes $r \times (n-r)$ and $(n-r) \times (n-r)$ respectively.

Back
Full Screen
Close
Quit

Let $T\mid_U \in \mathrm{End}(U)$ denote the restriction of T to U. Show that A is the matrix of T with respect to the basis of U with which we started.

Exercise 8.2. Suppose we have a chain of T-invariant subspaces $U_1 \subset U_2 \cdots \subset U_t$, with $\dim U_i = d_i$. Start with a basis of U_1, expand it to a basis of U_2, expand again to a basis of U_3, etc. until finally we have a basis of V. What will the matrix of T look line in this basis?

8.2. Eigenvectors and Eigenvalues

The simplest invariant subspaces are the one-dimensional ones.

Definition 8.2. Let $T \in \mathrm{End}(V)$. Suppose there is a nonzero vector $v \in V$ and $\lambda \in F$ such that $Tv = \lambda v$. Then v is called an *eigenvector* of T with *eigenvalue* λ.

Theorem 8.1. *Let $T \in \mathrm{End}(V)$. The following are equivalent.*

1. *V has a basis consisting of eigenvectors of T.*
2. *There is a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}^{\mathcal{B}}$ is a diagonal matrix.*

Proof. Exercise. □

Definition 8.3. $T \in \mathrm{End}(V)$ is called *diagonalizable* if V has a basis of eigenvectors of T.

Definition 8.4. An $n \times n$ matrix is called diagonalizable if it is the matrix (in any basis) of a diagonalizable linear map. Thus a matrix is diagonalizable if and only if it is similar to a diagonal matrix.

Theorem 8.2. *Suppose $v_1, \ldots, v_k \in V$ are eigenvectors of $T \in \mathrm{End}(V)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Then $v_1, \ldots, v_k$ are linearly independent.*

Proof. Suppose for a contradiction that they are dependent. Choose a nontrivial dependence relation involving the smallest possible number (r, say) of the v_i. By changing notation if necessary, we can assume that the relation is

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = 0.$$

By minimality, we know that all the α_i are nonzero. Also it is clear that $r \geq 2$ (why?). Apply T and use the fact that we have eigenvectors to obtain

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_r \lambda_r v_r = 0.$$

Multiplying (8.2) by λ_1 and subtracting (8.2) yields

$$\alpha_2(\lambda_1 - \lambda_2)v_2 + \cdots + \alpha_r(\lambda_1 - \lambda_r)v_r = 0.$$

Since the λ_j are distinct, this is a nontrivial linear independence relation involving $r-1$ of the v_i, contradicting our assumption. $\square$

9. Quotient Spaces

Let V be a vector space over F and let U be a subspace of V. For each $v \in V$, let

$$v + U = \{v + u \mid u \in U\}.$$

These are certain subsets of V.

Lemma 9.1. *Let v, $v' \in V$. Then $v + U = v' + U$ if and only if $v - v' \in U$.*

We define a relation on V by the rule that $v \sim v'$ iff $v + U = v' + U$. This is an equivalence relation, as is easily checked. Therefore, V is partitioned into the sets $v + U$. It is clear that $v \in v + U$ but it is important to remember that v does not have any special status among the elements of $v + U$ since if v' is another element of $v + U$ then $v' + U = v + U$.

Definition 9.1. The set whose elements are the distinct sets $v + U$ is denoted by V/U.

Theorem 9.2. *V/U is a vector space under the addition and scalar multiplication given by the following formulae:*

1. $(v + U) + (v' + U) := (v + v') + U \ \forall v,\ v' \in V$.
2. $\alpha(v + U) := \alpha v + U \ \forall v \in V,\ \forall \alpha \in F$.

Proof. Since we are attempting to define addition of equivalence classes in terms of addition of representative elements in the classes, we must check that our definition is independent of our choice of $v \in v+U$ and $v' \in v'+U$. Suppose $v_1+U = v+U$ and $v_1'+U = v'+U$. Then by Lemma 9.1, we have $v_1-v \in U$ and $v_1'-v' \in U$. Then since U is a subspace, we have $(v_1+v_1')-(v+v') \in U$. Therefore $(v_1+v_1')+U = (v+v')+U$.

This shows that addition is well-defined. Similarly, one can check that the definition of scalar mutiplication does depend on choice of representative element in the class $v+U$. It is now very easy to check that the axioms of a vector space hold for V/U. Because of the way we have defined the addition and scalar multiplication for V/U, you will see that the validity of the axioms for V/U will follow from their validity for V. □

Definition 9.2. The map $\pi : V \to V/U$ defined by $\pi v = v+U$ is called the natural map from V to V/U.

Exercise 9.1. Show that π is a linear map. What are its kernel and image?

Exercise 9.2. Suppose $T : V \to W$ is a linear map and the subspace U is contained in $\operatorname{Ker} T$. Prove that there is a linear map $\overline{T} : V/U \to W$ such that $\overline{T}(v+U) = Tv$, for all $v \in V$.

Definition 9.3. The map $\overline{T} : V/U \to W$ is called the linear map *induced by* T.

Exercise 9.3. Suppose $T \in \operatorname{End}(V)$ and U is a T-invariant subspace. Show that there exists a linear map $\overline{T} \in \operatorname{End}(V/U)$ such that $\overline{T}(v+U) = Tv+U$, for all $v \in V$.

Definition 9.4. The map $\overline{T} \in \operatorname{End}(V/U)$ is called the endomorphism of V/U *induced by* T.

Exercise 9.4. Suppose $T \in \operatorname{End}(V)$ and U is a T-invariant subspace. Let $v_1,\dots,$ v_k be a basis of U, and extend them to a basis $\mathcal{B}$: $v_1,\dots,$ v_n of V. Explain why $v_{k+1}+U,\dots,$ v_n+U form a basis of V/U. We have seen that the matrix of T in the basis $\mathcal{B}$ has the form

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}.$$

Let $T\mid_U \in \mathrm{End}(U)$ denote the restriction of T to U. In a previous exercise we saw that A ia the matrix of $T\mid_U$ with respect to the above basis of U. Now show that C is the matrix of $\overline{T}$ with respect to the above basis of V/U.

10. Spaces of Linear maps and the Dual Space

10.1. The space of linear maps

Definition 10.1. Let V and W be vector spaces over F. The set of all linear maps $T : V \to W$ is called $\mathrm{Hom}(V, W)$. Addition and scalar multiplication are defined by

$$(T + S)v = Tv + Sv, \quad (\alpha T)v = \alpha(Tv), \quad \forall T, S \in \mathrm{Hom}(V, W), \forall v \in V, \forall \alpha \in F.$$

Lemma 10.1. $\mathrm{Hom}(V, W)$ *is a vector space.*

Proof. Exercise. □

Exercise 10.1. Suppose $\dim V = n$ and $\dim W = m$ and let ordered bases $\mathcal{B}$ and $\mathcal{C}$ be given. Show that the mapping from $\mathrm{Hom}(V, W)$ to $\mathrm{Mat}_{m\times n}(F)$ sending a linear map T to its matrix $[T]^{\mathcal{C}}_{\mathcal{B}}$ is an isomorphism (a bijective linear map).

Exercise 10.2. Show that the set of all maps from any set X into a vector space V is a vector space using the formulae (10.1) for the operations. (Compare with Example 2.6.) Thus, $\mathrm{Hom}(V, W)$ could have been defined as a subspace of this larger space.

10.2. The dual space

Definition 10.2. The space $\mathrm{Hom}(V, F)$ is called the *dual space* of V and denoted V^*.

If V is finite dimensional, then we can already see from Exercise 10.1 that $\dim V^* = \dim V$, but let's take a closer look. Suppose $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis of V. For any n values $\mu_i \in F$, there exists a unique linear map sending v_i to μ_i

(What general principle is this?). Thus for each i we let x_i be the unique linear such that

$$x_i(v_j) = \begin{cases} 1, \text{if } i = j, \\ 0, \text{if } i \neq j, \end{cases}$$

Let $v \in V$, then $x_i(v)$ is equal to the i-th coordinate of v with respect to the basis $\mathcal{B}$. For this reason, the x_i are called the coordinate functions associated with the basis $\mathcal{B}$.

Lemma 10.2. *The set of the x_i form a basis of V^*.*

Proof. Exercise. □

The set of coordinate functions is also known as the *dual basis* of the basis $\mathcal{B}$.

Definition 10.3. Let A be an $m \times n$ matrix. The *transpose* of A, denoted A^t is the $n \times m$ matrix whose i-th column is the i-th row of A.

Exercise 10.3. If the matrix product AB is defined, then $(AB)^t = B^t A^t$.

Notation 10.4. In situations where one must consider both a V and V^* at the same time, it is convenient to write the coordinate vectors of V with respect to $\mathcal{B}$ are as columns and the coordinate vectors of V^* with respect to the dual basis $\mathcal{B}^*$ in transposed form as rows. For example, we have for each $f \in V^*$ and $v \in V$, the formula

$$f(v) = [\,f\,]_{\mathcal{B}^*}{}^t [\,v\,]_{\mathcal{B}},$$

where we regard the 1×1 matrix on the right hand side as a field element.

Remark 10.1. Exercise 10.3 shows that when we use row vectors as coordinate vectors, the matrix of a linear map should be transposed and multiply on the right.

Definition 10.5. Let $T : V \to W$ be a linear map. We define $T^* : W^* \to V^*$, by $(T^* f)v = f(Tv)$, $\forall f \in W^*$ and $\forall v \in V$.

Lemma 10.3. *Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V and W respectively and let $\mathcal{B}^*$ and $\mathcal{C}^*$ be the dual bases. Let $T \in \mathrm{Hom}(V, W)$ and let $A = [\,T\,]_{\mathcal{B}}^{\mathcal{C}}$. Then the following hold.*

1. *If we write the coordinate vectors of V^* in the usual way as column vectors, then $[\,T^*\,]_{\mathcal{C}^*}^{\mathcal{B}^*} = A^t$.*

2. *For $v \in V$, $f \in W^*$, we have*

$$(T^* f)(v) = f(Tv) = [\,f\,]_{\mathcal{C}^*}{}^t A [\,v\,]_{\mathcal{B}}$$

Proof. The proof of this lemma is problem 1 on assignment 3. (Hint: The notation and remark above are used to prove (2.) □

Next we consider the double dual $V^{**} := (V^*)^*$. Let V be any vector space (possibly infinite-dimensional) For each $v \in V$, the mapping $E_v : V^* \to F$ is defined by $E_v(f) = f(v)$, $\forall f \in V^*$.

Exercise 10.4. Check that E_v is linear, therefore an element of V^{**}.

Next, define a map $\Psi : V \to V^{**}$ by $\Psi(v) = E_v$.

Theorem 10.4. *The map Ψ is an injective linear map. If V is finite-dimensional, then Ψ is an isomorphism.*

Proof. This is exercise 2 of assignment 3. □

Remark 10.2. We have already seen that any two vector spaces over F of the same dimension are isomorphic, hence V is isomorphic with V^{**} and even with

V^*. So why do we bother with Theorem 10.4? The answer to this question is that the map Ψ has been defined without making any choices of bases and is therefore "canonical" or "natural". Another example was the map $T^* : W^* \to V^*$ and another is the natural map from V to V/U. Can you think of any others?

Exercise 10.5. Find a natural isomorphism of $\mathrm{Hom}(F, V)$ with V.

Exercise 10.6. For subspace U of a vector space V, let

$$U^\perp = \{f \in V^* \mid f(u) = 0, \forall u \in U\}.$$

Show that $U^\perp$ is a subspace of V^*. Prove that $\dim U + \dim U^\perp = \dim V$. (Hint: consider the linear map from V^* to U^* induced by the inclusion map of U into V.)

11. Direct Sums

Definition 11.1. If U_1 and U_2 are subspaces of V, their sum $U_1 + U_2$ is defined to be the subspace generated by $U_1 \cup U_2$. More generally, for any collection $\{U_i \mid i \in I\}$ of subspaces, the sum is the subspace they generate.

Exercise 11.1. Let U and W be subspaces of V. Prove that $\dim U + \dim W = \dim(U + W) + \dim(U \cap W)$.

Definition 11.2. If U_1 and U_2 are subspaces of V, we say that V is the direct sum of U_1 and U_2 if $V = U_1 + U_2$ and $U_1 \cap U_2 = \{0\}$. In this case we write $V = V_1 \oplus V_2$.

More generally, we say that V is the direct sum of subspaces $U_1, \ldots, U_r$ if $V = U_1 + \cdots + U_r$ and for each i, $U_i \cap (\sum j \neq iU_j) = \{0\}$. In this case we write $V = U_1 \oplus \cdots \oplus U_r$.

These definitions should be compared to the definition of linear independence.

Lemma 11.1. *The following are equivalent:*

1. $V = U_1 \oplus \cdots \oplus U_r$.
2. *Every element* $v \in V$ *has a unique expression as* $v = u_1 + \cdots + u_r$*, with* $u_i \in U_i$.

Proof. Suppose $V = U_1 \oplus \cdots \oplus U_r$. Then certainly each element can be written as in (2). Suppose $v = u_1 + \cdots + u_r = u'_1 + \cdots + u'_r$. Then for any i, we have $u'_i - u_i = \sum_{j \neq i}(u_j - u'_j)$. This element therefore belongs to both U_i and to $\sum_{j \neq i} U_j$, so must be the zero vector. Thus $u_i = u'_i$. Since i was arbitrary, we see that the expression for v is unique. Thus, (1) implies (2).

Now assume (2). It is then clear that $V = U_1 + \cdots U_r$, so we must show that the second condition in the definition of direct sum holds. Let i be arbitrary. Suppose we have $v \in U_i \cap (\sum_{j\neq i} U_j)$. Then there exist elements $u_j \in U_j$ ($j \neq i$ such that $v = \sum_{j\neq i} u_j$. Then $-v + \sum_{j\neq i} u_j = \mathbf{0}$ and, since $v \in U_i$, this is an expression for $\mathbf{0}$ as in (2). Since $\mathbf{0}$ is also equal the sum of the $\mathbf{0}$ vectors from each subspace, the uniqueness in (2) implies $v = \mathbf{0}$. Since i was arbnitrary, we have proved that $V = U_1 \oplus \cdots \oplus U_r$. $\square$

Lemma 11.2. *Suppose $V = U_1 \oplus \cdots \oplus U_r$. Let $\mathcal{B}_i$ be a basis of U_i. Then $\mathcal{B} := \cup_{i=1}^r \mathcal{B}_i$ is a basis of V.*

Proof. Exercise. $\square$

Exercise 11.2. Show that if U_1 is a subspace of V, then there exists a subspace U_2 such that $V = U_1 \oplus U_2$. Explain with examples, why there may be many different possiblities for the subspace U_2.

Exercise 11.3. Suppose $V = U_1 \oplus \cdots \oplus U_r$ and each U_i is a T-invariant subspace, for $T \in \mathrm{End}(V)$. Show that V has a basis in which the matrix of T has "block-diagonal" form, consisting of r square blocks down the main diagonal and zeroes elsewhere.

12. Minimal polynomial

Let $T \in \mathrm{End}(V)$. We have defined how to multiply endomorphisms (by composition) and also how to add endomorphisms. Therefore it is clear what is meant by a polynomial in T (the constant term is the corresponding scalar multiple of id_V). Since composition and addition of endomorphisms satisfy the distributive law, it also makes sense to factorize a polynomial in T. Let $p(x) \in F[x]$ be a polynomial. We say that T *satisfies* $p(x)$ if the endomorphism $p(T)$ is the zero endomorphism.

Lemma 12.1. *If* $\dim V = n$ *then every endomorphism* T *satisfies a polynomial of degree* n^2.

Proof. The endomorphisms $\mathrm{id}_V, T, \ldots T^{n^2}$ must be linearly dependent since $\dim \mathrm{End}(V) = n^2$. A linear dependence relation gives the desired polynomial realtion for T. □

Definition 12.1. We say that T has minimal polynomial $m(x)$ if $m(x)$ has highest coefficient 1 and has the smallest degree of any nonzero polynomial satisfied by T.

Lemma 12.2. *If* T *has minimal polynomial* $m(x)$ *and* T *satisfies* $p(x)$, *then* $m(x)$ *divides* $p(x)$. *In particular,* T *has a unique minimal polynomial.*

Proof. By long division of polynomials, we can write

$$p(x) = q(x)m(x) + r(x),$$

where $q(x)$ and $r(x)$ are polynomials and $r(x)$ is either zero or of degree strictly less than the degree of $m(x)$. Then,

$$0 = p(T) = q(T)m(T) + r(T) = r(T).$$

Thus, T satisfies $r(x)$. By minimality of $m(x)$, we must have $r(x) = 0$. □

Corollary 12.3. *Let $T \in \mathrm{End}(V)$ and let U be a T-invariant subspace of V. Let $T\mid_U \in \mathrm{End}(U)$ be the restriction of T to U and let $\overline{T} \in \mathrm{End}(V/U)$ be the induced endomorphism of V/U. The the minimal polynomials of $T\mid_U$ and $\overline{T}$ divide the minimal polynomial of T.*

Proof. It follows from the definitions that $T\mid_U$ and $\overline{T}$ satisfy any polynomial which T satisfies. So the result follows from Lemma 12.2. □

Theorem 12.4. *T is diagonalizable if and only if its minimal polynomial factorizes as a product of distinct linear factors.*

Proof. Suppose first that T is diagonalizable. Then it has a basis $\mathcal{B}$ of eigenvectors. Let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues which occur. Then since the elements of $\mathcal{B}$ which have eigenvalue λ_i are mapped to $\mathbf{0}$ by $T - \lambda_i \mathrm{id}_V$, it is clear that all elements of $\mathcal{B}$ are mapped to $\mathbf{0}$ by $\prod_{i=1}^r (T - \lambda_i \mathrm{id})$, so by Lemma 12.2, the minimal polynomial of T factorizes into distinct linear factors. We now prove the converse. Let $m(x) = \prod_{i=1}^r (x - \lambda_i)$ be the minimal polynomial of T. Set $m_1(x) = \prod_{i=1}^{r-1} (x - \lambda_i)$. By minimality of $m(x)$, we that there exists $v \in V$ with $m_1(T)v \neq \mathbf{0}$. Since

$$\mathbf{0} = m(T)v = [(T - \lambda_r \mathrm{id}_V) m_1(T)]v = (T - \lambda_r \mathrm{id}_V)(m_1(T)v),$$

we see that $\mathrm{Ker}\,(T - \lambda_r \mathrm{id}_V) \neq \{\mathbf{0}\}$. Set $V_{\lambda_r} = \mathrm{Ker}\,(T - \lambda_r \mathrm{id}_V)$ and set $U = \mathrm{Im}\,(T - \lambda_r \mathrm{id}_V)$. Then by problem 1 on the second assignment, we know that V_{λ_r} and U are T-invariant subspaces. We will show that

$$V = V_{\lambda_r} \oplus U.$$

Suppose we know this. Then $T\mid_U$ satisfies $m(x)$ (by the corollary above) and $\dim U < \dim V$. Therefore if we argue by induction on dimension, the inductive hypothesis would tell us that U has a basis of eigenvectors. Since $T\mid_{V_{\lambda_r}}$ certainly has a basis of eigenvectors, it follows that V has a basis of eigenvectors, i.e T is diagonalizable.

It remains to prove (12). By considering dimensions, we see that it is enough to show $V_{\lambda_r} \cap U = \{\mathbf{0}\}$ (Why?). Let $v \in V_{\lambda_r} \cap U$. Then since $v \in V_{\lambda_r}$, we have $Tv = \lambda_r v$, so $m_1(T)v = [\prod_{i=1}^{r-1}(\lambda_r - \lambda_i)]v$, a nozero scalar multiple of v. On the other hand, since $v \in U$ there is some $v' \in V$ with $v = (T - \lambda_r \mathrm{id}_V)v'$, so $m_1(T)v = m(T)v' = \mathbf{0}$. This proves (12). $\square$

Exercise 12.1. Suppose T and S are two diagonalizable endomorphisms of V such that $ST = TS$. Show that they are simultaneously diagonalizable, that is, there is a basis consisting of eigenvectors for both T and S.

Exercise 12.2. Prove that if T is diagonalizable with distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, then the minimum polynomial of T is $\prod_{i=1}^{r}(x - \lambda_i)$.

Exercise 12.3. Show (using only what we have proved) that every eigenvalue of an endomorphism T is a root of the minimal polynomial of T. Conversely, show that each root in F of the minimal polynomial of T is an eigenvalue of T.

Definition 12.2. Let $T \in \mathrm{End}(V)$ and $\lambda \in \mathrm{F}$. The subspace

$$V_\lambda = \{v \in V \mid Tv = \lambda v\}$$

is called the *eigenspace* of T with eigenvalue λ, or λ-eigenspace for short.

Note that $V_\lambda = \{v \in V \mid (T - \lambda \mathrm{id})v = \mathbf{0}\}$. This description motivates the following definition.

Definition 12.3. Let $T \in \mathrm{End}(V)$ and $\lambda \in \mathrm{F}$. The subspace

$$E_\lambda = \{v \in V \mid \exists k \in \mathbb{N}, (T - \lambda \mathrm{id})^k v = \mathbf{0}\}$$

is called the *algebraic eigenspace* of T with eigenvalue λ. Note that the k may be different for different v.

Thus, V_λ consists of those vectors sent to $\mathbf{0}$ by $T-\lambda\mathrm{id}$, while E_λ consists of those vectors sent to $\mathbf{0}$ by some power of $T-\lambda\mathrm{id}$. Clearly $V_\lambda \subseteq E_\lambda$. If you think about it, you have already given examples where the inclusion is proper. Sometimes, for extra precision, V_λ is called the *geometric eigenspace*. From now on, for simplicity, we will write $T-\lambda$ to mean $T-\lambda\mathrm{id}_V$.

Theorem 12.5. *Suppose all roots of the minimal polynomial of T (i.e. all the eigenvalues of T) lie in F. Then V is the direct sum of the algebraic eigenpaces of T.*

Proof. The proof of this theorem is very similar to that of Theorem 12.4. We argue by induction on the dimension of V. The theorem is evident for dimension zero or 1. So let $\dim V > 1$ and assume that the theorem holds for all vector spaces of smaller dimension. We will try to use these hypotheses to deduce the theorem for V. The hypothesis says that we may factorize the minimal polynomial of T as

$$m(x) = \prod_{i=1}^{r} (x-\lambda_i)^{e_i}, \quad \text{where } e_i \in \mathbb{N}.$$

Set $m_1(x) = \prod_{i=1}^{r-1}(x-\lambda_i)^{e_i}$. Then, by minimality of $m(x)$, there exists $v \in V$ such that $m_1(T)v \neq \mathbf{0}$. Since

$$\mathbf{0} = m(T)v = (T-\lambda_r)^{e_r}(m_1(T)v),$$

we see that the endomorphism $S = (T-\lambda_r)^{e_r}$ has a nonzero kernel. Since S commutes with T, we know that $\operatorname{Ker} S$ and $\operatorname{Im} S$ are T-invariant subspaces. We claim that

$$V = \operatorname{Ker} S \oplus \operatorname{Im} S.$$

Suppose we can prove this claim. Then by Corollary 12.3, we know that the minimal polynomial of the restriction of T to Im S also has all its roots in F. (Indeed its roots are among those of $m(x)$.) Since Im S has smaller dimension than V, th inductive hypothesis tells us that Im S is the direct sum of algebraic eigenspaces of (the restriction of) T. Since Ker S is an algebraic eigenspace for the restriction of T, it then follows easily that V is a direct sum of algebraic eigenspaces of T, so the theorem is proved. The proof of (12) is similar to the corresponding proof of Theorem 12.4 and is left as an exercise. □

This theorem reduced the study of endomorphisms whose eigenvalues lie in F, to the study of their restrictions to the algebraic eigenspaces. So we can consider the situation $V = E_\lambda$. If we are interested in T, we may as well consider $T-\lambda$. We know that $(T-\lambda)^e = 0$ for some $e \geq 1$. Endomorphisms which become zero when raised to some power are called *nilpotent*. You looked at these in Assignment 2. In summary, our discussion reduces the question of classifying all endomorphisms whose eigenvalues all lie in F to the question of classifying all nilpotent endomorphisms.

Remark 12.1. A field F is *algebraically closed* if every non-constant polynomial has a root, or, equivalently, if every non-constant polynomial is a product of linear factors. The field of complex numbers is an example of an algebraically closed field (Fundamental Theorem of Algebra). It can be proved that every field is (isomorphic with) a subfield of an algebraically closed field. Clearly if F is algebraically closed, then all vector space endomorphisms satisfy the hypotheses of Theorem 12.5

Exercise 12.4. Show that if an endomorphism has all its eigenvalues in F, then there is a basis in which its matrix is in triangular form. (This is Problem 4 on Assignment 4.)

13. Bilinear Forms

Definition 13.1. A *bilinear form* on a vector space V is a map $B : V \times V \to F$ (of two variables) which is linear in each variable:

$$B(\lambda u + v, w) = \lambda B(u, w) + B(v, w), \forall u, v, w \in V, \lambda \in F$$

and

$$B(w, \lambda u + v) = \lambda B(w, u) + B(w, u), \forall u, v, w \in V, \lambda \in F$$

Example 13.1. The dot product of $\mathbb{R}^n$.

Example 13.2. The usual inner product $\mathbf{a}.\mathbf{b} = \sum a_i \bar{b}_i$ of $\mathbb{C}^n$ is not bilinear. It is an example of a *sesquilinear* or *hermitian* form, which is discussed later.

Example 13.3. The cross product of $\mathbb{R}^3$ is not an example of a bilinear form, since it maps into V, not F, though it does satisfy the bilnearity properties.

Example 13.4. The function $B((a, b), (c, d)) = ad - bc$ is a bilinear form on F^2.

Exercise 13.5. Let B be a bilinear form on V. Fix $v \in V$ and define $f_v(w) = B(v, w)$ and $g_v(w) = B(w, v) \forall w \in V$. Show that f_v and g_v lie in V^* and that the mappings $\theta_L : v \mapsto f_v$ and $\theta_R : v \mapsto g_v$ are linear maps from V to V^*.

13.1. Coordinates

Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Then because of bilinearity, a bilinearform B is determined completely by the n^2 values $B(v_i, v_j)$. Let A be the matrix whose (i, j) entry is $B(v_i, v_j)$. Then one can check that

$$B(v, w) = [\, v \,]_{\mathcal{B}}^t A [\, w \,]_{\mathcal{B}}.$$

Conversely, starting with any $n \times n$ matrix A, the equation (13.1) can be used to define a function B which, by the rules of matrix algebra, is a bilinear form with matrix A in the given basis. Thus, once a basis is fixed, we have a bijection between bilinear forms and $n \times n$ matrices.

We now want to see what happens to the matrix of B if we change basis. Let $\mathcal{B}'$ be a second basis. Then, for all $v \in V$, we have $[\,v\,]_{\mathcal{B}} = P[\,v\,]_{\mathcal{B}'}$, where $P = [\,\mathrm{id}\,]^{\mathcal{B}}_{\mathcal{B}'}$ is the invertible matrix for the change of basis. Substituting in (13.1) gives

$$B(v, w) = [\,v\,]^t_{\mathcal{B}'} P^t A P [\,w\,]_{\mathcal{B}}.$$

It follows that the matrix of B with respect to $\mathcal{B}'$ is $P^t A P$.

Definition 13.2. Two $n \times n$ matrices A and C (over F) are said to be *congruent* if there exists an invertible $n \times n$ matrix P such that $P^t A P = C$.

It is easily verified that congruence is an equivalence relation. We have seen that the matrices of a bilinear form B with respect to different bases are congruent. Conversely, congruent matrices represent the same bilinear form in different bases, since given a bases $\mathcal{B}$ every invertible $n \times n$ matrix is of the form $[\,\mathrm{id}\,]^{\mathcal{B}}_{\mathcal{B}'}$ for a unique basis $\mathcal{B}'$.

Exercise 13.6. Suppose B has matrix A in the basis $\mathcal{B}$. Let $\mathcal{B}^*$ be the dual basis How is A related to $[\,\theta_L\,]^{\mathcal{B}^*}_{\mathcal{B}}$ and $[\,\theta_R\,]^{\mathcal{B}^*}_{\mathcal{B}}$? Deduce that the kernels of θ_L and θ_R have the same dimension.

Definition 13.3. Let B be a bilinear form. The *left radical* of B, denoted $\mathrm{Rad}_L(B)$ is the set

$$\{v \in V \mid B(v, w) = 0, \forall w \in V\}.$$

It is clear that $\mathrm{Rad}_L(B)$ is a subspace – in fact it is nothing other than the kernel of θ_L. The *right radical* $\mathrm{Rad}_R(B)$ is defined analogously. By Exercise 13.6 the left radical is zero if and only if the right radical is zero.

Definition 13.4. If the left and right radicals of a bilinear form are zero, then the form is said to be *non-singular*.

Thus, if B is nonsingular then θ_L and θ_R are isomorphisms of V onto V^* and conversely if either map is an isomorphism then B is nonsingular.

13.2. Symmetric bilinear forms

We assume in this subsection that the characteristic of F is not 2 i.e. $1 + 1 \neq 0$ in F.

Definition 13.5. A bilinear form B on V is *symmetric* if $B(v, w) = B(w, v)$, $\forall v$, $w \in V$.

Definition 13.6. Given a symmetric bilinear form B, the function $Q : V \to F$ defined by $Q(v) = \frac{1}{2}B(v, v)$ is called the quadratic form associated with B.

The reason for this terminology is that if $x_1, \ldots, x_n$ are coordinate functions for V, then by (13.1) the function Q is expressed as a homogeneous quadratic polynomial in the x_i.

Lemma 13.1. *(Polarization) We have for v, $w \in V$,*

$$2B(v, w) = B(v + w, v + w) - B(v, v) - B(w, w).$$

The symmetric bilinear form B is completely determined by the quadratic form Q.

Proof. Direct calculation. □

Let us now consider the matrix of a symmetric bilinear form. For any given basis, the matrix A of B will be *symmetric*, i.e. $A = A^t$. Our next task is to look for bases in which this matrix has a nice form.

Theorem 13.2. *Assume that the characteristic of F is not 2. Then V has a basis $\{v_i\}$ such that $B(v_i, v_j) = 0$ for $i \neq j$.*

Proof. We argue by induction on $\dim V$. There is nothing to prove if $\dim V = 1$. Also, the theorem holds if B is identically zero. So we assume that $\dim V = n > 1$ and that $B(v, w) \neq 0$ for some v, $w \in V$. By polarization, there must exist $v_1 \in V$ with $B(v_1, v_1) \neq 0$. Set $U = \{v \in V \mid B(v_1, v) = 0\}$.Then U is equal to the kernel of the map $f_{v_1} \in V^*$. This map is nonzero , hence its image is F. It follows that $\dim U = n - 1$. By induction, U has a basis $v_2, \ldots v_n$ such that $B(v_i, v_j) = 0$ for all i, $j = 2, \ldots, n$ with $i \neq j$. By definition of U, this holds also if we include v_1. It remains to check that $v_1, \ldots, v_n$ form a basis for V and for this, it suffuces to show that they span V. Let $v \in V$. Then it is easy to check that $v - \frac{B(v,v_1)}{B(v_1,v_1)} v_1 \in U$. □

Corollary 13.3. *If the characteristic of F is not 2, then every symmetric matrix is congruent to a diagonal matrix.*

The number of nonzero diagonal entries is called the *rank* of B (or of the associated quadratic form). It is simply the rank (in the usual sense) of any matrix of B, since this is not changed under congruence.

Which diagonal matrices are congruent? The answer is usually very difficult, but over the complex and real numbers we can give the complete answer.

Corollary 13.4. *Every symmetric complex matrix is congruent to a diagonal matrix with 1s and 0s on the diagonal.*

Proof. Exercise □

Corollary 13.5. *Every quadratic form on a vector space over $\mathbb{C}$ (or any algebraically closed field of chracteristic $\neq 2$) can be written in suitable coordinates as $Q = x_1^2 + \cdots x_r^2$, where r is the rank of Q.*

Corollary 13.6. *Every symmetric real matrix is congruent to a diagonal matrix with* 1*s,* -1*s and* 0*s on the diagonal.*

Proof. Exercise. □

Corollary 13.7. *Every quadratic form on a real vector space can be written in suitable coordinates as* $Q = x_1^2 + \cdots x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$.

Exercise 13.7. Prove that if a real quadratic form is expressed as in Corollary 13.7 in two coordinate systems, then the number of 1s and -1s and 0s will be the same in each coordinate system. (This result is called *Sylvester's Law of Inertia*). The number $p - q$ is called the *signature*.

Definition 13.7. A quadratic form on a real vector space is *positive definite (resp. semi-definite)* if $Q(v) > 0$ (resp. $q(v) \geq 0$) for every nonzero vector v. If $-Q$ is positive definite (semi-definite) we say Q is *negative definite (semidefinite)*.

We have shown that in suitable coordinates, all positive definite forms look just like the usual squared length function of euclidean geometry. Another form of importance in geometry and physics is the form $Q = -x_1^2 - x_2^2 - x_3^3 + x_4^2$, which defines the metric on spacetime in the theory of Special Relativity. This form is *definite*, neither positive nor negative (semi-)definite.

Exercise 13.8. Show that there are no positive definite quadratic forms on a complex vector space.

Exercise 13.9. Find the rank and signature of the form $x_1x_2 + x_2x_3 + x_3x_1$.

Exercise 13.10. Show that if a quadratic form Q on a real vector space is indefinite, then there is a nonzero vector v such that $Q(v)$=0. Give an example to show that this is false for vector spaces over the rational numbers.

Notation 13.8. In the case of symmetric bilinear forms, there is no distinction between θ_L and θ_R, $\mathrm{Rad}_L(B)$ and $\mathrm{Rad}_R(B)$, so we can just use θ, and refer to *the* radical $\mathrm{Rad}(B)$.

13.3. Skew-symmetric bilinear forms

A bilinear form is called *skew-symmetric* or or *antisymmetric* or *alternating* if $B(v, v) = 0$ for every $v \in V$. For such a form we have

$$0 = B(v+u, v+u) = B(v, v) + B(v, u) + B(u, v) + B(u, u) = B(v, u) + B(u, v).$$

so $B(u, v) = -B(v, u)$. From this it follows that the left and right radicals are equal.

Let U be a subspace of V such that $V = U \oplus \mathrm{Rad}(B)$. If we form a basis of V from bases of U and of $\mathrm{Rad}(B)$, then the matrix of B in this basis will have the form

$$\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix},$$

where A_1 is the matrix of the restriction of B to U, with respect to the given basis of U. It follows from this that the restriction of B to U is non-singular. In this way, the study of skew-symmetric forms can be reduced to that of non-singular ones. The same remark is also valid for symmetric bilinear forms.

Theorem 13.8. *Let B be a non-singular skew-symmetric bilinear form on V. Then V has even dimension $n = 2m$ and there is a basis $e_1, \ldots, e_m, f_1, \ldots, f_m$ such that $B(e_i, f_j) = \delta_{ij}$, $B(e_i, e_j) = 0$ and $B(f_i, f_j) = 0$ for all i, j.*

Proof. We argue by induction on dimension. It is clear that the only skew-symmetric form on a one-dimensional vector space is the zero form. Suppose $\dim V \geq 2$. Let

$e \in V$ be any nonzero vector. By nonsingularity and bilinearity, there exists $f \in V$ such that $B(e, f) = 1$. Let $V_1 = \langle e, f\rangle$ and let $U = \{u \in V | B(u, v) = 0 \forall v \in V_1\}$. Then U is a subspace of V (equal to the intersection of $\operatorname{Ker} \theta(e)$ and $\operatorname{Ker} \theta(f)$). We claim that $V = V_1 \oplus U$. Let $v \in V$. Then $v - B(v, f)e - B(e, v)f \in U$, so $V = V_1 + U$. Next suppose $u \in U \cap V_1$. Then $B(u, w) = 0$ for all elements w of V_1, since $u \in U$, and for all elements of U, since $w \in V_1$. But since $V = V_1 + U$, this shows that $u = \mathbf{0}$. Thus we have proved our claim that $V = V_1 \oplus U$. Next we observe that the restriction of the form to U is nonsingular. This is because if $u \in U$ satisfies $B(u, u')$ for all $u' \in U$, then since $V = V_1 \oplus U$, we would have $B(u, v) = 0$ for all $v \in V$, so $u = \mathbf{0}$.

Let $n = \dim V$. Then $\dim U = n - 2$ and U has a nonsingular alternating bilinear form. Thew inductive hypothesis applies to tell us first that $n - 2$ is even. Hence n is even. Write $n = 2m$. The inductive hypothesis also tells us that U has a basis and a basis $e_2, \ldots, e_m$,$f_2, \ldots, f_m$ such that $B(e_i, f_j) = \delta_{ij}$, $B(e_i, e_j) = 0$ and $B(f_i, f_j) = 0$ for all i, j. Setting $e_1 = e$ and $f_1 = f$, it is now easy to verify that we have the required basis for V. □

A basis of the type described in this theorem is called a *symplectic basis*.

This theorem tells us that all skew-symmetric forms on V which have the same rank will look alike in suitable coordinates, or, in terms of matrices, any two $n \times n$ skew-symmetric matrices of the same rank are congruent.

14. Hermitian Forms

Throughout this section, F is taken to be the field $\mathbb{C}$ of complex numbers. Since there are no positive definite quadratic forms over the complex numbers, another type of form is needed to define the notion of distance in complex spaces.

Definition 14.1. A *hermitian form* on a complex vector space V is a mapping $h : V \times V \to \mathbb{C}$ which is linear in the first variable and satisfies *hermitian symmetry* $h(v, w) = \overline{h(w, v)}$, where the bar denotes complex conjugation.

Thus, h is"conjugate-linear" in the second variable.

The theory parallels the discussion of symmetric bilinear forms.

As in the case of symmetric bilinear forms we have a *polarization* identity. Set $f(v) = h(v, v)$. Then

$$4h(v, w) = f(v + w) - f(v - w) + if(v + iw) - if(v - iw).$$

so h and f determine each other.

Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Let $a_{ij} = h(v_i, v_j)$. Then h is determined by the matrix $A = (a_{ij})$ and this matrix satisfies the relation $A = \overline{A}^t$. Any matrix satisfying this relation is called *hermitian*. The form can be computed by the formula

$$h(v, w) = [\,v\,]^t_{\mathcal{B}} A \overline{[\,w\,]_{\mathcal{B}}}.$$

If we change basis from $\mathcal{B}$ to $\mathcal{B}'$, the matrix of the form will be changed to $P^t A \overline{P}$, where $P = [\,\mathrm{id}\,]^{\mathcal{B}}_{\mathcal{B}'}$.

Theorem 14.1. *Let h be a hermitian form on a complex vector space. Then in suitable coordinates, we can write*

$$h(x, x) = |x_1|^2 + \cdots + |x_p|^2 - |x_{p+1}|^2 - \cdots - |x_r|^2,$$

where p and r depend only on h and not on the choice of basis.

Proof. Exercise. If you need help, look at the case of quadratic forms. □

We can define positive (or negative) (semi-)definite hermitian forms just as for quadratic forms. We see that the hermitian forms where $p = r = n$ are positive definite.

15. Euclidean and Unitary Spaces

15.1. Euclidean spaces and the orthogonal group

Definition 15.1. A euclidean space is a vector space over $\mathbb{R}$ which has a positive definite symmetric bilinear form $\langle , \rangle$.

The reduction theorem for real quadratic forms (Cor. 13.6) tells us that such a form always has an orthonormal basis.

Let V be a euclidean space of dimension n.

Definition 15.2. For a subspace $W \subseteq V$, we set $W^\perp = \{u \in V \mid \langle u, w\rangle = 0 \forall w \in W\}$.

Exercise 15.1. $W^\perp$ is a subspace of V and $V = W \oplus W^\perp$.

Definition 15.3. Let H be a hyperplane of V, i.e a subspace of dimension $n-1$. Then $H^\perp$ is one-dimensional and there exists a unit vector $u \in H^\perp$, determined up to sign. The *reflection* in H τ_H is the unique linear transformation which is the identity on H and sends u to $-u$. One can check that the explicit formula is

$$\tau_H(v) = v - 2\langle v, u\rangle u,$$

by verifying that the linear map thus given has the desired effect on H and u.

We next consider those endomorphisms of V which preserve the form. Such an endomorphism must have kernel zero, since the form is positive definite, hence is invertible, and it is easily seen that the inverse also preserves the form. Thus, the totality of such endomorphisms forms a group.

Definition 15.4. The group of linear maps of V which preserve the given positive definite quadratic form is called the *orthogonal group* and denoted $O(n)$ or $O_n(\mathbb{R})$. Its elements are called orthogonal transformations.

It is easily checked that reflections in hyperplanes are orthogonal transformations. The next theorem shows that these reflections generate the whole orthogonal group.

Exercise 15.2. Let A be the matrix of an orthogonal transformation with respect to an orthonormal basis. Show that $A^tA = I$. (Such a matrix is called an orthogonal matrix.)

Theorem 15.1. *Every orthogonal transformation is a product of $\leq n$ reflections in hyperplanes.*

Proof. Let $\sigma \in O(n)$. We argue by induction on n, the case $n = 1$ being trivial, since then the only orthogonal transformations are the identity and multiplication by -1, which is a reflection in the hyperplane $\{\mathbf{0}\}$. So we assume $n > 1$ and that the theorem holds in dimension $n-1$. Assume first that there exists nonzero vector v with $\sigma v = v$. Set $U = \langle v\rangle^\perp$. Then $V = \langle v\rangle \oplus U$, where U is of dimension $n-1$. Next, U is σ-invariant, for if $u \in U$ we have

$$\langle \sigma u, v\rangle = \langle \sigma u, \sigma v\rangle = \langle u, v\rangle = 0.$$

Let τ be the restriction of σ to U. Then τ is an orthogonal trnasformation of U. By induction,

$$\tau = \tau_1 \cdots \tau_r,$$

a product of r reflections τ_i in hyperplanes K_i of H. where $r \leq n-1$.

Let $H_i = K_i + \langle v\rangle$. Then H_i is a hyperplane of V. Let σ_i be the endomorphism of V which is the identity on $< v >$ and τ_i on H. Then one can check that σ_i is the reflection in the hyperplane H_i of V. Thus, we obtain

$$\sigma = \sigma_1 \cdots \sigma_r$$

as a product of $r \leq n-1$ reflections. This was all under the assumption that σ fixes the vector v. Now we treat the general case. Suppose $x \in V$ and $\sigma x \neq x$. Set $u = \sigma x - x$ and let $H = < u >^{\perp}$. We have

$$< \sigma x + x, \sigma x - x > = < \sigma x, \sigma x > - < \sigma x, x > - < x, \sigma x > - < x, x > = 0.$$

so $\sigma x + x \in H$. Let τ be the reflection in H. Then $\tau(\sigma x + x) = \sigma x + x$ and $\tau(\sigma x - x) = x - \sigma x$. Adding, we see $\tau \sigma x = x$. Thus, the transformation $\tau\sigma$ fixes a vector and by that case, we can write

$$\tau\sigma = \sigma_1 \cdots \sigma_r$$

as a porduct of $r \leq n-1$ reflections. Finally, since $\tau^{-1} = \tau$, we have

$$\sigma = \tau\sigma_1 \cdots \sigma_r,$$

a product of at most n reflections. □

15.2. Unitary spaces and the unitary group

Definition 15.5. A vector space over $\mathbb{C}$ which has a positive definite hermitian form $<,>$ is called a *unitary space.*

Let V be an n-dimensional unitary space. By Assignment 5, V has an orthonormal basis.

We next consider the transformations of V which preserve the form, called unitary transformations. The same reasoning as for euclidean spaces shows that the set of such transformations form a group, called the *unitary group* and denoted $U(n)$ or $U_n(\mathbb{C})$.

Back
Full Screen
Close
Quit

Exercise 15.3. Let A be the matrix of a unitary transformation with respect to an orthonormal basis. Show that $\overline{A}^t A = I$. (Such matrices are called unitary matrices).

Note that the unitary matrices which have real entries are precisely the orthogonal matrices. Thus general facts proved about unitary matrices also apply to orthogonal matrices.

Theorem 15.2. *A unitary transformation is diagonalizable and each eigenvalue λ satisfies $|\lambda| = 1$*

Proof. We argue by induction on n. If $n = 1$, it is easily seen that a unitary transformation is a multiplication by λ_1 with $|\lambda| = 1$. So assume $n > 1$ and that the theorem holds for $n-1$. Let $T \in U(n)$. Since $\mathbb{C}$ is algebraically closed, T has an eigenvector v_1, and as in the case $n = 1$, the eigenvalue λ_1 for v_1 satisfies $|\lambda| = 1$. Let $W =< v_1 >^{\perp}$. Then $V =< v_1 > \oplus W$ and W is a T-invariant subspace of dimension $n-1$, with the restriction of T to W a unitary transformation of W (cf. the proof of the previous theorem for orthogonal transformations). Therefore, by induction, W has a basis $v_2, \ldots, v_n$ of eigenvectors of T with eigenvalues $\lambda_2, \ldots \lambda_n$ satisfying $|\lambda_i| = 1$. The basis $v_1, v_2, \ldots v_n$ is the required basis of V of eigenvectors for T, completing the proof of the theorem. □

Corollary 15.3.

1. *The minimal polynomial of a unitary transformation factors into distinct linear factors of the form $(x - e^{i\theta})$.*
2. *The minimal polynomial of an orthogonal transformation of a euclidean space has the following possible irreducible factors, each with multiplicity one:linear factors of the form $(x \pm 1)$ and quadratic factors of the form $x^2 - 2\cos\theta x + 1$.*

Proof. Part (1) is immediate from the theorem. To prove (2), we remember that an orthogonal matrix is just unitary matrix which has real entries. Therefore, its minimal polynomial is a real polynomial, whose complex roots are as in (1). Because the polynomial is real, the roots other than ± 1 must occur in complex conjugate pairs $e^{i\theta}$ and $e^{-i\theta}$. Each such pair gives an irredicble quadratic factor $(x - e^{i\theta})(x - e^{-i\theta}) = x^2 - 2\cos\theta + 1$ of the minimal polynomial. $\square$

16. Self-Adjoint Linear Maps

The discussion of this section for unitary spaces can easily be modified for euclidean spaces, but for clarity, we will consider just the unitary case.

Let V be a unitary space and $<,>$ its positive definite hermitian form.

Definition 16.1. An endomorphism T of V is *self-adjoint* if $< Tv, w > = < v, Tw >$ for all v, $w \in V$.

The following results about a self-adjoint map T are simple consequences of the definition and the proofs are left as exercises.

Lemma 16.1. *The eigenvalues of T are real.*

Lemma 16.2. *Eigenvectors of T corresponding to distinct eigenvalues are orthogonal, i.e the value of the form is zero on such as pair.*

Let V_λ be the (geometric) eigenspace of T for the eigenvalue λ.

Lemma 16.3. *We have a decomposition*

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r},$$

into mutually orthogonal eigenspaces.

Proof. We argue by induction on dimension. Let λ_1 be an eigenvalue of T. Then, as for any subspace, we can write

$$V = V_{\lambda_1} \oplus (V_{\lambda_1})^{\perp}.$$

It is clear that V_{λ_1} is T-invariant. Let $w \in (V_{\lambda_1})^{\perp}$. Then for $v \in V_{\lambda_1}$, we have

$$< Tw, v > = < w, Tv > = < w, \lambda_1 v > = 0,$$

which shows that $(V_{\lambda_1})^\perp$ is T-invariant. The inductive hypothesis applies to $(V_{\lambda_1})^\perp$, so that

$$(V_{\lambda_1})^\perp = V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r},$$

where $\lambda_2,\ldots,\lambda_r$ are the eigenvalues of T on $(V_{\lambda_1})^\perp$. Note that none of these is equal to λ_1 since if $w \in (V_{\lambda_1})^\perp$ was an eigenvector with eigenvalue λ_1, it would be orthogonal to itself. Thus we have the desired decmposition in to eigenspaces, which are mutually orthogonal by the previous lemma. □

We can now state and prove the *Spectral Theorem.*

Theorem 16.4. *Assume $T \in \operatorname{End}(V)$ is self-adjoint. Then V has an orthonormal basis of eigenvectors for T.*

Proof. Simply choose an orthonormal basis for each eigenspace V_{λ_i} of the preceding lemma. □

Exercise 16.1. Show that the matrix of a self-adjoint linear map with respect to an orthonormal basis is hermitian, i.e $A = \overline{A}^t$. Show also that any hermitian matrix arises in this way.

Corollary 16.5. *Let A be an hermitian matrix. Then there is a unitary matrix U such that*

$$UAU^{-1} = UA\overline{U}^t$$

is diagonal.

Since we have seen also that hermitian matrices arise as matrices of hermitian forms, the above corollary also yields the following result about hermitian forms.

Corollary 16.6. *Let h be an hermitian form. Then there exists an orthonormal basis in which the form is diagonalized.*

17. Notation

- $\mathbb{N}$ the natural numbers 0, 1, 2,...
- $\mathbb{Z}$ the integers
- $\mathbb{Q}$ the rational numbers
- $\mathbb{R}$ the real numbers
- $\mathbb{C}$ the complex numbers
- $\forall$ universal quantifier ("for all")
- $\exists$ existential quantifier ("there exists")
- iff logical equivalence ("if and only if")
- $\mathbf{0}$ the zero vector of a vector space
- $[v]_{\mathcal{B}}$ the coordinate vector of v with respect to the ordered basis $\mathcal{B}$
- $\mathrm{Mat}_{m\times n}(F)$ the set of $m \times n$ matrices with entries in F
- $\mathrm{Poly}(F)$ the set of polynomials with coefficients in F
- $\phi_{\mathcal{B}}$ the coordinate map with respect to the ordered basis $\mathcal{B}$
- $[T]_{\mathcal{B}}^{\mathcal{C}}$ the matrix of a linear map with respect to ordered bases $\mathcal{B}$ and $\mathcal{C}$
- id, id_V the identity map of V, $\mathrm{id}(v) = v$
- I_n the identity $n \times n$ matrix