A Second Semester of Linear Algebra

S. E. Payne

February 25, 2004

Contents

1 Preliminaries **7**

1.1 Fields 8
1.2 Groups 9
1.3 Matrix Algebra 9
1.4 Linear Equations Solved with Matrix Algebra 11

2 Vector Spaces **13**

2.1 Definition of Vector Space 13
2.1.1 Prototypical Example 13
2.1.2 A Second Example 14
2.2 Basic Properties of Vector Spaces 14
2.3 Subspaces 15
2.4 Sums and Direct Sums of Subspaces 16
2.5 Exercises 17

3 Dimensional Vector Spaces **19**

3.1 The Span of a List 19
3.2 Linear Independence and the Concept of Basis 20
3.3 Exercises 25

4 Linear Transformations **27**

4.1 Definitions and Examples 27
4.2 Kernels and Images 28
4.3 Rank and Nullity Applied to Matrices 30
4.4 Projections 31
4.5 Bases and Coordinate Matrices 31
4.6 Matrices as Linear Transformations 33

4.7 Change of Basis 36
4.8 Exercises 39

5 Polynomials **41**
5.1 Algebras 41
5.2 The Algebra of Polynomials 43
5.3 Lagrange Interpolation 45
5.4 Polynomial Ideals 47
5.5 Exercises 51

6 Determinants **55**
6.1 Determinant Functions 55
6.1.3 n-Linear Alternating Functions 58
6.1.5 A Determinant Function - The Laplace Expansion 59
6.2 Permutations & Uniqueness of Determinants 61
6.2.1 A Formula for the Determinant 61
6.3 Additional Properties of Determinants 66
6.3.1 If A is a unit in $M_n(K)$, then $\det(A)$ is a unit in K. . . 66
6.3.2 Triangular Matrices 66
6.3.3 Transposes 66
6.3.4 Elementary Row Operations 67
6.3.5 Triangular Block Form 67
6.3.6 The Classical Adjoint and the Laplace Expansion . . . 68
6.3.8 Characteristic Polynomial of a Linear Map 71
6.3.9 Coefficients of the Characteristic Polynomial 71
6.3.11 The Companion Matrix of a Polynomial 73
6.3.13 The Cayley-Hamilton Theorem 74
6.3.14 Cramer's Rule 76
6.4 Deeper Results with Some Applications* 76
6.4.1 Block Matrices whose Blocks Commute* 77
6.4.2 Tensor Products of Matrices* 78
6.4.3 The Cauchy-Binet Theorem-A Special Version* 79
6.4.5 The Matrix-Tree Theorem* 81
6.4.10 The Cauchy-Binet Theorem - A General Version* . . . 83
6.4.12 The General Laplace Expansion* 85
6.4.14 Determinants, Ranks and Linear Equations* 86
6.5 Exercises 93

7 Operators and Invariant Subspaces **95**
7.1 Eigenvalues and Eigenvectors 95
7.2 Upper-Triangular Matrices 97
7.3 Invariant Subspaces of Real Vector Spaces 101
7.4 Two Commuting Linear Operators* 102
7.5 Commuting Families of Operators* 106
7.6 The Fundamental Theorem of Algebra* 109
7.7 Exercises 113

8 Inner Product Spaces **115**
8.1 Inner Products 115
8.2 Orthonormal Bases 124
8.3 Orthogonal Projection and Minimization 126
8.4 Linear Functionals and Adjoints 130
8.5 The Rayleigh Principle* 132
8.6 Exercises 135

9 Operators on Inner Product Spaces **139**
9.1 Self-Adjoint Operators 139
9.2 Normal Operators 143
9.3 Decomposition of Real Normal Operators 146
9.4 Positive Operators 149
9.5 Isometries 151
9.6 The Polar Decomposition 155
9.7 The Singular-Value Decompostition 157
9.7.3 Two Examples 161
9.8 Pseudoinverses and Least Squares* 164
9.9 Norms, Distance and More on Least Squares* 169
9.10 Exercises 175

10 Decomposition WRT a Linear Operator **177**
10.1 Powers of Operators 177
10.2 The Algebraic Multiplicity of an Eigenvalue 179
10.3 Elementary Operations 182
10.4 Transforming Nilpotent Matrices 184
10.5 A "Jordan Form" for Real Matrices 187
10.6 Exercises 194

11 Matrix Functions* **199**
11.1 Operator Norms and Matrix Norms* 199
11.2 Polynomials in an Elementary Jordan Matrix* 202
11.3 Scalar Functions of a Matrix* 204
11.4 Scalar Functions as Polynomials* 208
11.5 Power Series* 210
11.6 Commuting Matrices* 216
11.7 A Matrix Differential Equation* 221
11.8 Exercises 223

12 Infinite Dimensional Vector Spaces* **225**
12.1 Partially Ordered Sets & Zorn's Lemma* 225
12.2 Bases for Vector Spaces* 226
12.3 A Theorem of Philip Hall* 227
12.4 A Theorem of Marshall Hall, Jr.* 230
12.5 Exercises* 232

Chapter 1

Preliminaries

Preface to the Student

This book is intended to be used as a text for a second semester of linear algebra either at the senior or first-year-graduate level. It is written for you under the assumption that you already have successfully completed a first course in linear algebra and a first course in abstract algebra. The first short chapter is a very quick review of the basic material with which you are supposed to be familiar. If this material looks new, this text is probably not written for you. On the other hand, if you made it into graduate school, you must have already acquired some background in modern algebra. Perhaps all you need is to spend a little time with your undergraduate texts reviewing the most basic facts about equivalence relations, groups, matrix computations, row reduction techniques, and the basic concepts of linear independence, span, and basis in the context of $\mathcal{R}^n$.

On the other hand, some of you will be ready for a more advanced approach to the material covered here than we can justify making a part of the course. For this reason I have included some starred sections that may be skipped without disturbing the general flow of ideas, but that might be of interest to some students. If material from a starred section is ever cited in a later section, that section will necessarily also be starred.

For the material we do cover in detail, we hope that you will find our presentation to be thorough and our proofs to be complete and clear. However, when we indicate that you should verify something for yourself, that means you should use paper and pen and write out the appropriate steps in detail.

One major difference between your undergraduate linear algebra text and

this one is that we discuss abstract vector spaces over arbitrary fields instead of restricting our discussion to the standard space of n-tuples of real numbers. A second difference is that the emphasis is on linear transformations from one vector space over the field F to a second one, rather than on matrices over F. However, a great deal of the general theory can be effortlessly translated into results about matrices.

1.1 Fields

The fields of most concern in this text are the complex numbers $\mathcal{C}$, the real numbers $\mathcal{R}$, the rational numbers $\mathcal{Q}$, and the finite Galois fields $GF(q)$, where q is a prime power. In fact, it is possible to work through the entire book and only use only the fields $\mathcal{R}$ and $\mathcal{C}$. And if finite fields are being considered, most the time it is possible to use just those finite fields with a prime number of elements, i.e., the fields $\mathcal{Z}_p \sim \mathcal{Z}/p\mathcal{Z}$, where p is a prime integer, with the algebraic operations just being addition and multiplication modulo p. For the purposes of reading this book it is sufficient to be able to work with the fields just mentioned. However, we urge you to pick up your undergraduate abstract algebra text and review what a field is. For most purposes the symbol F denotes an arbitrary field except where inner products and/or norms are involved.

Kronecker delta When the underlying field F is understood, the symbol δ_{ij} (called the *Kronecker delta* denotes the element $1 \in F$ if $i = j$ and the element $0 \in F$ if $i \neq j$. Occasionally it means 0 or 1 in some other structure, but the context should make that clear.

If you are not really familiar with the field of complex numbers, you should spend a little time getting acquainted. A complex number $\alpha = a + bi$, where $a, b \in \mathcal{R}$ and $i^2 = -1$, has a conjugate $\overline{\alpha} = a - bi$, and $\alpha \cdot \overline{\alpha} = a^2 + b^2 = |\alpha|^2$. It is easily checked that $\overline{\alpha \cdot \beta} = \overline{\alpha} \cdot \overline{\beta}$. Note that $|\alpha| = 0$ if and only if $\alpha = 0$. You should show how to compute the multiplicative inverse of any nonzero complex number.

We define the square-root symbol $\sqrt{\ }$ as usual: For $0 \leq x \in \mathcal{R}$, put $\sqrt{x} = |y|$ where y is a real number such that $y^2 = x$.

The following two lemmas are often useful.

Lemma 1.1.1. *Every complex number has a square root.*

Proof. Let $\alpha = a + bi$ be any complex number. With the definition of $\sqrt{}$ given above, and with $\gamma = \sqrt{a^2 + b^2} \geq |a|$, we find that

$$\left(\sqrt{\frac{\gamma + a}{2}} \pm i \sqrt{\frac{\gamma - a}{2}} \right)^2 = a \pm |b| i.$$

Now just pick the sign $\pm$ so that $\pm |b| = b$. □

Lemma 1.1.2. *Every polynomial of odd degree with real coefficients has a (real) zero.*

Proof. It suffices to prove that a monic polynomial

$$P(x) = x^n + a_1 x^{n-1} + \cdots + a_n$$

with some $a_i \neq 0$, with $a_1, \ldots, a_n \in \mathcal{R}$ and n odd has a zero. Put $a = |a_1| + |a_2| + \cdots + |a_n| + 1 > 1$, and $\epsilon = \pm 1$. Then $|a_1(\epsilon a)^{n-1} + \cdots + a_{n-1}(\epsilon a) + a_n| \leq |a_1|(a)^{n-1} + \cdots + |a_{n-1}|a + |a_n| \leq (a-1)(a^{n-1}) < a^n$. It readily follows that $P(a) > 0$ and $P(-a) < 0$. Hence by the Intermediate Value Theorem there is a $\lambda \in (-a, a)$ such that $P(\lambda) = 0$. □

1.2 Groups

Let G be an arbitrary set and let $\circ : G \times G \to G$ be a binary operation on G. Usually we denote the image of $(g_1, g_2) \in G \times G$ under the map $\circ$ by $g_1 \circ g_2$. Then you should know what it means for $(G, \circ)$ to be a group. If this is the case, G is an abelian group provided $g_1 \circ g_2 = g_2 \circ g_1$ for all $g_1, g_2 \in G$. Our primary example of a group will be a vector space whose elements (called vectors) form an abelian group under vector addition. It is also helpful if you remember how to construct the quotient group G/N, where N is a normal subgroup of G. However, this latter concept will be introduced in detail in the special case where it is needed.

1.3 Matrix Algebra

An $m \times n$ *matrix* A over F is an m by n array of elements from the field F. We may think of A as an ordered list of m row vectors from F^n or equally

well as an ordered list of n column vectors from F^m. The element in row i and column j is usually denoted A_{ij}. The symbol $M_{m,n}(F)$ denotes the set of all $m \times n$ matrices over F. It is readily made into a vector space over F. For $A, B \in M_{m,n}(F)$ define $A + B$ by

$$(A + B)_{i,j} = A_{ij} + B_{ij}.$$

Similarly, define scalar multiplication by

$$(aA)_{ij} = aA_{ij}.$$

Just to practice rehearsing the axioms for a vector space, you should show that $M_{m,n}(F)$ really is a vector space over F. (See Section 2.1.)

If A and B are $m \times n$ and $n \times p$ over F, respectively, then the product AB is an $m \times p$ matrix over F defined by

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}.$$

Lemma 1.3.1. *Matrix multiplication, when defined, is associative.*

Proof. (Sketch) If A, B, C are $m \times n$, $n \times p$ and $p \times q$ matrices over F, respectively, then

$$((AB)C)_{ij} = \sum_{l=1}^{p}(AB_{il})C_{lj} = \sum_{l=1}^{p}(\sum_{k=1}^{n} A_{ik}B_{kl})C_{lj} =$$

$$= \sum_{1 \le k \le n; 1 \le l \le p} A_{ik}B_{kl}C_{lj} = (A(BC))_{ij}.$$

□

The following observations are not especially deep, but they come in so handy that you should think about them until they are second nature to you.

Obs. 1.3.2. *The ith row of AB is the ith row of A times the matrix B.*

Obs. 1.3.3. *The jth column of AB is the matrix A times the jth column of B.*

If A is $m \times n$, then the $n \times m$ matrix whose (i, j) entry is the (j, i) of A is called the *transpose* of A and is denoted A^T.

Obs. 1.3.4. *If A is $m \times n$ and B is $n \times p$, then $(AB)^T = B^T A^T$.*

Obs. 1.3.5. *If A is $m \times n$ with columns $C_1, \ldots, C_n$ and $X = (x_1, \ldots, x_n)^T$ is $n \times 1$, thenAX is the linear combination of columns of A given by $\sum_{j=1}^{n} x_j C_j$.*

Obs. 1.3.6. *If A is $m \times n$ with rows $\alpha_1, \ldots, \alpha_m$ and $X = (x_1, \ldots, x_m)$, then XA is the linear combination of rows of A given by $\sum_{i=1}^{m} x_i \alpha_i$.*

At this point you should review block multiplication of matrices.

1.4 Linear Equations Solved with Matrix Algebra

For each i, $1 \leq i \leq m$, let

$$\sum_{j=1}^{n} a_{ij} x_j = b_j$$

be a linear equation in the indeterminates $x_1, \ldots, x_n$ with the coefficients a_{ij} and b_j all being real numbers. One of the first things you learned to do in your undergraduate linear algebra course was to replace this system of linear equations with a single matrix equation of the form

$$A\vec{x} = \vec{b}, \text{ where } A = (a_{ij}) \text{ is } m \times n \text{ with } m \text{ rows and } n \text{ columns,}$$

$$\text{and } \vec{x} = (x_1, \ldots, x_n)^T, \quad \vec{b} = (b_1, \ldots, b_m)^T.$$

You then augmented the matrix A with the column $\vec{b}$ to obtain

$$A' = \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ a_{21} & \ldots & a_{2n} & b_2 \\ \vdots & \ldots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array} \right).$$

At this point you performed elementary row operations on the matrix A' so as to replace the submatrix A with a row-reduced echelon matrix R. You then used this matrix $(R \mid \vec{b'})$ to read off all sorts of information about the original matrix A and the system of linear equations. The matrix R has r

nonzero rows for some r, $0 \leq r \leq m$. The leftmost nonzero entry in each row of R is a 1, and it is the only nonzero entry in its column. Then the matrix $(R \,|\, \vec{b'})$ is used to solve the original system of equations by writing each of the "leading variables" as a linear combination of the other (free) variables. The r nonzero rows of R form a basis for the row space of A, the vector subspace of $\mathcal{R}^n$ spanned by the rows of A. The columns of A in the same positions as the leading 1's of R form a basis for the column space of A, i.e., the subspace of $\mathcal{R}^m$ spanned by the columns of A. There are $n-r$ free variables. Let each of these variables take a turn being equal to 1 while the other free variables are all equal to 0, and solve for the other leading variables using the matrix equation $R\vec{x} = \vec{0}$. This gives a basis of the (right) nullspace of A, i.e., a basis of the space of solutions to the system of homogeneous equations obtained by replacing $\vec{b}$ with $\vec{0}$. In particular we note that r is the dimension of both the row space of A and of the column space of A. We call r the *rank of* A. The dimension of the right null space of A is $n-r$. Replacing A with its transpose A^T, so that the left null space of A becomes the right null space of A^T, shows that the left null space of A has dimension $m-r$.

A good reference for the basic material is the following text: David C. Lay, LINEAR ALGEBRA AND ITS APPLICATIONS, 3rd Edition, Addison Wesley, 2003.

Chapter 2

Vector Spaces

Linear algebra is primarily the study of linear maps on finite-dimensional vector spaces. In this chapter we define the concept of vector space and discuss its elementary properties.

2.1 Definition of Vector Space

A *vector space over the field* F is a set V together with a binary operation "+" on V such that $(V, +)$ is an abelian group, along with a *scalar multiplication* on V (i.e., a map $F \times V \to V$) such that the following properties hold:

1. For each $a \in F$ and each $v \in V$, av is a unique element of V with $1\ v = v$ for all $v \in V$. Here 1 denotes the multiplicative identity of F.

2. Scalar multiplication distributes over vector addition: $a(u + v) = (au) + (av)$, which is usually written as $au + av$, for all $a \in F$ and all $u, v \in V$.

3. $(a + b)u = au + bu$, for all $a, b \in F$ and all $u \in V$.

4. $a(bv) = (ab)v$, for all $a, b \in F$ and all $v \in V$.

2.1.1 Prototypical Example

Let S be any nonempty set and let F be any field. Put $V = F^S = \{f : S \to F : f \text{ is a function}\}$. Then we can make V into a vector space over F as follows. For $f, g \in V$, define the vector sum $f + g : S \to F$ by

$$(f + g)(s) = f(s) + g(s) \text{ for all } s \in S.$$

Then define a scalar multiplication $F \times V \to V$ as follows:

$$(af)(s) = a(f(s)) \text{ for all } a \in F,\ f \in F,\ s \in S.$$

It is a very easy but worthwhile exercise to show that with this vector addition and this scalar multiplication, V is a vector space over F. It is also interesting to see that this family of examples includes (in some abstract sense) all the examples of vector spaces you might have studied in your first linear algebra course. For example, let $F = \mathcal{R}$ and let $S = \{1, 2, \ldots, n\}$. Then each $f : \{1, 2, \ldots, n\} \to \mathcal{R}$ is given by the n-tuple $(f(1), f(2), \ldots, f(n))$. So with almost no change in your point of view you can see that this example is essentially just $\mathcal{R}^n$ as you knew it before.

2.1.2 A Second Example

Let F be any field and let x be an indeterminate over F. Then the ring $F[x]$ of all polynomials in x with coefficients in F is a vector space over F. In Chapter 4 we will see how to view $F[x]$ as a subspace of the special case of the preceding example where $S = \{0, 1, 2, \ldots\}$. However, in this case, any two elements of $F[x]$ can be multiplied to give another element of $F[x]$, and $F[x]$ has the structure of a commutative ring with 1. It is even an integral domain! In fact, it is a *linear algebra*. See the beginning of Chapter 4 for the definition of a linear algebra, a term that really ought to be defined somewhere in a Linear Algebra course. (Look up any of these words that you are unsure about in your abstract algebra text.) (Note: Our convention is that the zero polynomial has degree equal to $-\infty$.)

If we fix the nonnegative integer n, then

$$\mathcal{P}_n = \{f(x) \in F[x] : degree(f(x)) \leq n\}$$

is a vector space with the usual addition of polynomials and scalar multiplication.

2.2 Basic Properties of Vector Spaces

We are careful to distinguish between the zero 0 of the field F and the zero vector $\vec{0}$ in the vector space V.

Theorem 2.2.1. *(Properties of zero) For $a \in F$ and $v \in V$ we have the following:*

$$av = \vec{0} \text{ if and only if } a = 0 \in F \text{ or } v = \vec{0} \in V.$$

Proof. First suppose that $a = 0$. Then $0v = (0+0)v = 0v + 0v$. Now adding the additive inverse of $0v$ to both sides of this equation yields $\vec{0} = 0v$, for all $v \in V$. Similarly, if $v = \vec{0}$, we have $a\vec{0} = a(\vec{0} + \vec{0}) = a\vec{0} + a\vec{0}$. Now add the additive inverse of $a\vec{0}$ to both sides to obtain $\vec{0} = a\vec{0}$.

What remains to be proved is that if $av = \vec{0}$, then either $a = 0$ or $v = \vec{0}$. So suppose $av = \vec{0}$. If $a = 0$ we are done. So suppose $a \neq 0$. In this case a has a multiplicative inverse $a^{-1} \in F$. So $v = (a^{-1}a)v = a^{-1}(av) = a^{-1}\vec{0} = \vec{0}$ by the preceding paragraph. This completes the proof. □

Let $-v$ denote the additive inverse of v in V, and let -1 denote the additive inverse of $1 \in F$.

Lemma 2.2.2. *For all vectors $v \in V$, $(-1)v = -v$.*

Proof. The idea is to show that if $(-1)v$ is added to v, the the result is $\vec{0}$, the additive identity of V. So, $(-1)v + v = (-1)v + 1v = (-1+1)v = 0v = \vec{0}$ by the preceding result. □

2.3 Subspaces

Definition A subset U of V is called a *subspace* of V provided that with the vector addition and scalar multiplication of V restricted to U, the set U becomes a vector space in its own right.

Theorem 2.3.1. *The subset U of V is a subspace of V if and only if the following properties hold:*

(i) $U \neq \emptyset$.

(ii) *If $u, v \in U$, then $u + v \in U$.*

(iii) *If $a \in F$ and $u \in U$, then $au \in U$.*

Proof. Clearly the three properties all hold if U is a subspace. So now suppose the three properties hold. Then U is not empty, so it has some vector u. Then $0u = \vec{0} \in U$. For any $v \in U$, $(-1)v = -v \in U$. By these properties and

property (ii) of the Theorem, $(U, +)$ is a subgroup of $(V, +)$. (Recall this from your abstract algebra course.) It is now easy to see that U, with the addition and scalar multiplication inherited from V, must be a vector space, since all the other properties hold for U automatically because they hold for V. □

2.4 Sums and Direct Sums of Subspaces

Definition Let $U_1, \ldots, U_m$ be subspaces of V. The *sum* $U_1 + U_2 + \cdots + U_m$ is defined to be

$$\sum_{i=1}^{m} U_i := \{u_1 + u_2 + \cdots + u_m : u_i \in U_i \text{ for } 1 \leq i \leq m\}.$$

You are asked in the exercises to show that the sum of subspaces is a subspace.

Definition The set $\{U_1, \ldots, U_m\}$ of subspaces is said to be *independent* provided that if $\vec{0} = u_1 + u_2 + \cdots + u_m$ with $u_i \in U_i$, $1 \leq i \leq m$, then $u_i = 0$ for all $i = 1, 2, \ldots, m$.

Theorem 2.4.1. *Let U_i be a subspace of V for $1 \leq i \leq m$. Each element $v \in \sum_{i=1}^{m} U_i$ has a* unique *expression of the form $v = \sum_{i=1}^{m} u_i$ with $u_i \in U_i$ for all i if and only if the set $\{U_1, \ldots, U_m\}$ is independent.*

Proof. Suppose the set $\{U_1, \ldots, U_m\}$ is independent. Then by definition $\vec{0}$ has a unique representation of the desired form (as a sum of zero vectors). Suppose that some $v = \sum_{i=1}^{m} u_i = \sum_{i=1}^{m} v_i$ has two such representations with $u_i,\ v_i \in U_i$, $1 \leq i \leq m$. Then $\vec{0} = v - v = \sum_{i=1}^{m}(u_i - v_i)$with $u_i - v_i \in U_i$ since U_i is a subspace. So $u_i = v_i$ for all i. Conversely, if each element of the sum has a unique representation as an element of the sum, then certainly $\vec{0}$ does also, implying that the set of subspaces is independent. □

Definition If each element of $\sum_{i=1}^{m} U_i$ has a unique representation as an element in the sum, then we say that the sum is the *direct sum* of the subspaces, and write

$$\sum_{i=1}^{m} U_i = U_1 \oplus U_2 \oplus \cdots \oplus U_m = \oplus \sum_{i=1}^{m} U_i.$$

Theorem 2.4.2. *Let $U_1, \ldots, U_m$ be subspaces of V for which $V = \sum_{i=1}^m U_i$. Then the following are equivalent:*

(i) $V = \oplus \sum_{i=1}^m U_i$.

(ii) $U_j \cap \sum_{1 \le i \le m; i \ne j} U_i = \{\vec{0}\}$ *for each* j, $1 \le j \le m$.

Proof. Suppose $V = \oplus \sum_{i=1}^m U_i$. If $u \in U_j \cap \sum_{1 \le i \le m; i \ne j} U_i$, say $u = -u_j \in U_j$ and $u = \sum_{1 \le i \le m; i \ne j} u_i$, then $\sum_{i=1}^m u_i = \vec{0}$, forcing all the u_i's equal to $\vec{0}$. This shows that (i) implies (ii). It is similarly easy to see that (ii) implies that $\vec{0}$ has a unique representation in $\sum_{i=1}^m U_i$. □

Note: When $m = 2$ this says that $U_1 + U_2 = U_1 \oplus U_2$ if and only if $U_1 \cap U_2 = \{\vec{0}\}$.

2.5 Exercises

1. Determine all possible subspaces of $F^2 = \{f : \{1, 2\} \to F : f \text{ is a function}\}$.

2. Prove that the intersection of any collection of subspaces of V is again a subspace of V.

3. Define the sum of a countably infinite number of subspaces of V and discuss what it should mean for such a collection of subspaces to be independent.

4. Prove that the set-theoretic union of two subspaces of V is a subspace if and only if one of the subspaces is contained in the other.

5. Prove or disprove: If U_1, U_2, W are subspaces of V for which $U_1 + W = U_2 + W$, then $U_1 = U_2$.

6. Prove or disprove: If U_1, U_2, W are subspaces of V for which $U_1 \oplus W = U_2 \oplus W$, then $U_1 = U_2$.

Chapter 3

Dimensional Vector Spaces

Our main concern in this course are the finite dimensional vector spaces, a concept that will be introduced in this chapter. However, in a starred section of the Appendix and with the help of the appropriate axioms from set theory we can show that every vector space has a well-defined dimension. The key concepts here are: span, linear independence, basis, dimension.

We assume throughout this chapter that V is a vector space over the field F.

3.1 The Span of a List

For the positive integer N let $\overline{N}$ be the ordered set $\overline{N} = \{1, 2, \ldots, N\}$, and let $\mathcal{N}$ be the ordered set $\{1, 2, \ldots, \}$ of all natural numbers. **Definition:** A *list* of elements of V is a function from some $\overline{N}$ to V or from $\mathcal{N}$ to V. Usually such a list is indicated by $(v_1, v_2, \ldots, v_m)$ for some positive integer m, or perhaps by $(v_1, v_2, \ldots)$ if the list is finite of unknown length or countably infinite. An important aspect of a list of vectors of V is that it is *ordered*. A second important difference between lists and sets is that elements of lists may be repeated, but in a set repetitions are not allowed. For most the work in this course the lists we consider will be finite, but it is important to keep an open mind about the infinite case.

Definition: Let $L = (v_1, v_2, \ldots)$ be a list of elements of V. We say that $v \in V$ is *in the span of* L provided there are finitely many scalars $a_1, a_2, \ldots, a_m \in F$ such that $v = \sum_{i=1}^{m} a_i v_i$. Then the set of all vectors in the span of L is said to *be* the *span of* L and is denoted $\text{span}(v_1, v_2, \ldots)$. By

convention we say that the span of the empty list ( ) is the zero space $\{\vec{0}\}$.

The proof of the following lemma is a routine exercise.

Lemma 3.1.1. *The span of any list of vectors of V is a subspace of V.*

If $\text{span}(v_1, \ldots, v_m) = V$, we say $(v_1, \ldots, v_m)$ *spans* V. A vector space is said to be *finite dimensional* if some finite list spans V. For example, let F^n denote the vector space of all column vectors with n entries from the field F, and let $\vec{e_i}$ be the column vector $(0, \ldots, 0,\ 1,\ 0 \ldots, 0)^T$ with $n-1$ entries equal to $0 \in F$ and a $1 \in F$ in position i. Then F^n is finite dimensional because the list $(e_1, e_2, \ldots, e_n)$ spans F^n.

Let $f(x) \in F[x]$ have the form $f(x) = a_0 + a_1 x + \cdots a_n x^n$ with $a_n \neq 0$. We say that $f(x)$ has *degree* n. The zero polynomial has degree $-\infty$. We let $\mathcal{P}_n(F)$ denote the set of all polynomials with degree at most n. Then $\mathcal{P}_n(F)$ is a subspace of $F[x]$ with spanning list $L = (1, x, x^2, \ldots, x^n)$.

3.2 Linear Independence and the Concept of Basis

One of the most fundamental concepts of linear algebra is that of linear independence.

Definition: A finite list $(v_1, \ldots, v_m)$ of vectors in V is said to be *linearly independent* provided the *only* choice of scalars $a_1, \ldots, a_m \in F$ for which $\sum_{i=1}^m a_i v_i = \vec{0}$ is $a_1 = a_2 = \cdots = a_m = 0$. An infinite list $(v_1, v_2, \ldots)$ of vectors in V is said to be *linearly independent* provided that for each positive integer m, the list $(v_1, \ldots, v_m)$ consisting of the first m vectors of the infinite list is linearly independent.

A finite *set* S of vectors of V is said to be *linearly independent* provided each list of distinct vectors of S (no repetition allowed) is linearly independent. An arbitrary set S of vectors of V is said to be linearly independent provided every finite subset of S is linearly independent.

Any subset of V or list of vectors in V is said to be *linearly dependent* provided it is not linearly independent. It follows immediately that a list $L = (v_1, v_2, \ldots)$ is linearly dependent provided there is some integer $m \geq 1$ for which there are m scalars $a_1, \ldots, a_m \in F$ such that $\sum_{i=1}^m a_i v_i = \vec{0}$ and not all the a_i's are equal to zero.

The following **Linear Dependence Lemma** will turn out to be extremely useful.

Lemma 3.2.1. *The nonempty list $L = (v_1, v_2, \ldots)$ of vectors in V with $v_1 \neq \vec{0}$ is linearly dependent if and only if there is some integer j such that the list $(v_1, \ldots, v_j)$ is a list consisting of the first j vectors of L for which the following hold:*

(i) $v_j \in span(v_1, \ldots, v_{j-1})$;

(ii) If the jth term v_j is removed from the list L, the span of the remaining list equals the span of L.

Proof. Suppose that the list $L = (v_1, v_2, \ldots)$ is linearly dependent and $v_1 \neq \vec{0}$. For some m there are scalars $a_1, \ldots, a_m$ for which $\sum_{i=1}^{m} a_i v_i = \vec{0}$ with at least one of the a_i not equal to 0. Since $v_1 \neq \vec{0}$, not all of the scalars $a_2, \ldots, a_m$ can equal 0. So let $j \geq 2$ be the largest index for which $a_j \neq 0$. Then

$$v_j = (-a_1 a_j^{-1})v_1 + (-a_2 a_j^{-1})v_2 + \cdots + (-a_{j-1} a_j^{-1})v_{j-1},$$

proving (i).

To see that (ii) also holds, just note that in any linear combination of vectors from L, if v_j appears, it can be replaced by its expression as a linear combination of the vectors $v_1, \ldots, v_{j-1}$.

For the converse, it should be immediately obvious that if (i) and (ii) hold, then the list L is linearly dependent. □

The following theorem is of major importance in the theory. It says that (finite) linearly independent lists are never longer than (finite) spanning lists.

Theorem 3.2.2. *Suppose that V is spanned by the finite list $L = (w_1, \ldots, w_n)$ and that $M = (v_1, \ldots, v_m)$ is a linearly independent list. Then $m \leq n$.*

Proof. We shall prove that $m \leq n$ by using an algorithm that is interesting in its own right. It amounts to starting with the list L and removing one w and adding one v at each step so as to maintain a spanning list. Since the list $L = (w_1, \ldots, w_n)$ spans V, adjoining any vector to the list produces a linearly dependent list. In particular,

$$(v_1, w_1, \ldots, w_n)$$

is linearly dependent with its first element different from $\vec{0}$. So by the Linear Dependence Lemma we may remove one of the w's so that the remaining list of length n is still a spanning list. Suppose this process has been carried out until a spanning list of the form

$$B = (v_1, \ldots, v_j, w'_1, \ldots, w'_{n-j})$$

has been obtained. If we now adjoin v_{j+1} to the list B by inserting it immediately after v_j, the resulting list will be linearly dependent. By the Linear Dependence Lemma, one of the vectors in this list must be a linear combination of the vectors preceding it in the list. Since the list $(v_1, \ldots, v_{j+1})$ must be linearly independent, this vector must be one of the w''s and not one of the v's. So we can remove this w' and obtain a new list of length n which still spans V and has $v_1, \ldots, v_{j+1}$ as its initial members. If at some step we had added a v and had no more w's to remove, we would have a contradiction, since the entire list of v's must be linearly independent. So we may continue the process until all the vectors $v_1, \ldots, v_m$ have been added to the list, i.e., $m \leq n$. □

Definition A vector space V over the field F is said to be *finite dimensional* provided there is a finite list that spans V.

Theorem 3.2.3. *If U is a subspace of the finite-dimensional space V, then U is finite dimensional.*

Proof. Suppose that V is spanned by a list of length m. If $U = \{\vec{0}\}$, then certainly U is spanned by a finite list and is finite dimensional. So suppose that U contains a nonzero vector v_1. If the list (v_1) does not span U, let v_2 be a vector in U that is not in span(v_1). By the Linear Dependence Lemma, the list (v_1, v_2) must be linearly independent. Continue this process. At each step, if the linearly independent list obtained does not span the space U, we can add one more vector of U to the list keeping it linearly independent. By the preceding theorem, this process has to stop before a linearly independent list of length $m + 1$ has been obtained, i.e., U is spanned by a list of length at most m, so is finite dimensional. □

Note: In the preceding proof, the spanning list obtained for U was also linearly independent. Moreover, we could have taken V as the subspace U on which to carry out the algorithm to obtain a linearly independent spanning list. This is a very important type of list.

Definition A *basis* for a vector space V over F is a list L of vectors of V that is a spanning list as well as a linearly independent list. We have just observed the following fact:

Lemma 3.2.4. *Each finite dimensional vector space V has a basis. By convention the empty set $\emptyset$ is a basis for the zero space $\{\vec{0}\}$.*

Lemma 3.2.5. *If V is a finite dimensional vector space, then there is a unique integer $n \geq 0$ such that each basis of V has length exactly n.*

Proof. If $B_1 = (v_1, \ldots, v_n)$ and $B_2 = (u_1, \ldots, u_m)$ are two bases of V, then since B_1 spans V and B_2 is linearly independent, $m \leq n$. Since B_1 is linearly independent and B_2 spans V, $n \leq m$. Hence $m = n$. □

Definition If the finite dimensional vector space V has a basis with length n, we say that V is *n-dimensional* or that *n is the dimension of V*. Moreover, each finite dimensional vector space has a well-defined dimension.

For this course it is completely satisfactory to consider bases only for finite dimensional spaces. However, students occasionally ask about the infinite dimensional case and we think it is pleasant to have a convenient treatment available. Hence we have included a treatment of the infinite dimensional case in an appendix that treats a number of special topics. For our general purposes, however, we are content merely to say that any vector space which is not finite dimensional is *infinite dimensional* (without trying to associate any specific infinite cardinal number with the dimension).

Lemma 3.2.6. *A list $\mathcal{B} = (v_1, \ldots, v_n)$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form*

$$v = a_1 v_1 + \cdots + a_n v_n. \tag{3.1}$$

Proof. First assume that $\mathcal{B}$ is a basis. Since it is a spanning set, each $v \in V$ can be written in the form given in Eq. 3.1. Since $\mathcal{B}$ is linearly independent, such an expression is easily seen to be unique. Conversely, suppose each $v \in V$ has a unique expression in the form of Eq. 3.1. The existence of the expression for each $v \in V$ implies that $\mathcal{B}$ is a spanning set. The uniqueness then implies that $\mathcal{B}$ is linearly independent. □

Theorem 3.2.7. *If $L = (v_1, \ldots, v_n)$ is a spanning list of V, then a basis of V can be obtained by deleting certain elements from L. Consequently every finite dimensional vector space has a basis.*

Proof. If L is linearly independent, it must already be a basis of V. If not, consider v_1. If $v_1 = \vec{0}$, discard v_1. If not, then leave L unchanged. Since L is assumed to be linearly dependent, there must be some j such that $v_j \in \text{span}(v_1, \ldots, v_{j-1})$. Choose the smallest j for which this is true and delete that v_j from L. This will yield a list L_1 of $n - 1$ elements that still

spans V. If L_1 is linearly independent, then L_1 is the desired basis of V. If not, then proceed as before to obtain a spanning list of size $n-2$. Continue this way, always producing spanning sets of shorter length, until eventually a linearly independent spanning set is obtained. □

Lemma 3.2.8. *Every linearly independent list of vectors in a finite dimensional vector space V can be extended to a basis of V.*

Proof. Let V be a finite dimensional space, say with $\dim(V) = n$. Let $L = (v_1, \ldots, v_k)$ be a linearly independent list. So $0 \leq k \leq n$. If $k < n$ we know that L cannot span the space, so there is a vector v_{k+1} not in $\text{span}(L)$. Then $L' = (v_1, \ldots, v_k, v_{k+1})$ must still be linearly independent. If $k = 1 < n$ we can repeat this process, adjoining vectors one at a time to produce longer linearly independent lists until a basis is obtained. As we have seen above, this must happen when we have an independent list of length n. □

Lemma 3.2.9. *If V is a finite-dimensional space and U is a subspace of V, then there is a subspace W of V such that $V = U \oplus W$. Moreover, $dim(U) \leq dim(V)$ with equality if and only if $U = V$.*

Proof. We have seen that W must also be finite-dimensional, so that it has a basis $\mathcal{B}_1 = (v_1, \ldots, v_k)$. Since $\mathcal{B}_1$ is a linearly independent set of vectors in a finite-dimensional space V, it can be completed to a basis $\mathcal{B} = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ of V. Put $W = \text{span}(v_{k+1}, \ldots, v_n)$. Clearly $V = U + W$, and $U \cap W = \{\vec{0}\}$. It follows that $V = U \oplus W$. The last part of the lemma should also be clear. □

Theorem 3.2.10. *Let $L = (v_1, \ldots, v_k)$ be a list of vectors of an n-dimensional space V. Then any two of the following properties imply the third one:*

(i) $k = n$;

(ii) L is linearly independent;

(iii) L spans V.

Proof. Assume that (i) and (ii) both hold. Then L can be completed to a basis, which must have exactly n vectors, i.e., L already must span V. If (ii) and (iii) both hold, L is a basis by definition and must have n elements by definition of dimension. If (iii) and (i) both hold, L can be restricted to form a basis, which must have n elements. Hence L must already be linearly independent. □

Theorem 3.2.11. *let U and W be finite-dimensional subspaces of the vector space V. Then*

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W).$$

Proof. Let $\mathcal{B}_1 = (v_1, \ldots, v_k)$ be a basis for $U \cap W$. Complete it to a basis $\mathcal{B}_2 = (v_1, \ldots, v_k, u_1, \ldots, u_r)$ of U and also complete it to a basis $\mathcal{B}_3 = (v_1, \ldots, w_1, \ldots, w_t)$ for W. Put $\mathcal{B} = (v_1, \ldots, v_k, u_1, \ldots, u_r, w_1, \ldots, w_t)$. We claim that $\mathcal{B}$ is a basis for $U + W$, from which the theorem follows.

First we show that $\mathcal{B}$ is linearly independent. So suppose that there are scalars $a_i,\ b_i, c_i \in F$ for which $\sum_{i=1}^{k} a_i v_i + \sum_{i=1}^{r} b_i u_i + \sum_{i=1}^{t} c_i w_i = \vec{0}$. It follows that $\sum_{i=1}^{k} a_i v_i + \sum_{i=1}^{r} b_i u_i = -\sum_{i=1}^{t} c_i w_i \in U \cap W$. Since $\mathcal{B}_2$ is linearly independent, this means all the b_i's are equal to 0. This forces $\sum_{i=1}^{k} a_i v_i + \sum_{i=1}^{t} c_i w_i = \vec{0}$. Since $\mathcal{B}_3$ is linearly independent, this forces all the a_i's and c_i's to be 0. But it should be quite clear that $\mathcal{B}$ spans $U + V$, so that in fact $\mathcal{B}$ is a basis for $U + W$. □

At this point the following theorem is easy to prove. We leave the proof as an exercise.

Theorem 3.2.12. *Let $U_1, \ldots, U_m$ be finite-dimensional subspaces of V with $V = U_1 + \cdots U_m$ and with $\mathcal{B}_i$ a basis for U_i, $1 \leq i \leq m$. The the following are equivalent:*

(i) $V = U_1 \oplus \cdots \oplus U_m$;
(ii) $dim(V) = dim(U_1) + dim(U_2) + \cdots + dim(U_m)$;
(iii) $(\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_m)$ is a basis for V.
(iv) The spaces $U_1, \ldots, U_m$ are linearly independent.

3.3 Exercises

1. Write out a proof of Lemma 3.1.1.

2. Let L be any list of vectors of V, and let S be the set of all vectors in L. Show that the intersection of all subspaces having S as a subset is just the span of L.

3. Show that any subset T of a linearly independent set S of vectors is also linearly independent, and observe that this is equivalent to the fact that if T is a linearly dependent subset of S, then S is also linearly dependent. Note: The empty set $\emptyset$ of vectors is linearly independent.

4. Show that the intersection of any family of linearly independent sets of vectors of V also linearly independent.

5. Give an example of two linearly dependent sets whose intersection is linearly independent.

6. Show that a list (v) of length 1 is linearly dependent if and only if $v = \vec{0}$.

7. Show that a list (v_1, v_2) of length 2 is linearly independent if and only if neither vector is a scalar times the other.

8. Let m be a positive integer. Let $V = \{f \in F[x] : \deg(f) = m \text{ or } f = \vec{0}\}$. Show that V is or is not a subspace of $F[x]$.

9. Prove or disprove: There is a basis of $\mathcal{P}_m(x)$ all of whose members have degree m.

10. Prove or disprove: there exists a basis (p_0, p_1, p_2, p_3) of $\mathcal{P}_3(F)$ such that

    (a) all the polynomials $p_0, \ldots, p_3$ have degree 3.

    (b) all the polynomials $p_0, \ldots, p_3$ give the value 0 when evaluated at 3.

    (c) all the polynomials $p_0, \ldots, p_3$ give the value 3 when evaluated at 0.

    (d) all the polynomials $p_0, \ldots, p_3$ give the value 3 when evaluated at 0 and give the value 1 when evaluated at 1.

11. Prove that if $U_1, U_2, \cdots, U_m$ are subspaces of V then $dim(U_1 + \cdots + U_m) \leq dim(U_1) + dim(U_2) + \cdots + dim(U_m)$.

12. Prove or give a counterexample: If U_1, U_2, U_3 are three subspaces of a finite dimensional vector space V, then

$$\dim(U_1 + U_2 + U_3) =$$

$$\dim(U_1) + \dim(U_2) + \dim(U_3)$$
$$-\dim(U_1 \cap U_2) - \dim(U_2 \cap U_3) - \dim(U_3 \cap U_1)$$
$$+\dim(U_1 \cap U_2 \cap U_3).$$

Chapter 4

Linear Transformations

4.1 Definitions and Examples

Throughout this chapter we let U, V and W be vector spaces over the field F.

Definition A function (or map) T from U to V is called a *linear map* or *linear transformation* provided T satisfies the following two properties:

(i) $T(u+v) = T(u) + T(v)$ for all $u, v \in U$.

and

(ii) $T(au) = aT(u)$ for all $a \in F$ and all $u \in U$.

These two properties can be combined into the following single property:

Obs. 4.1.1. *$T : U \to V$ is linear provided $T(au+bv) = aT(u)+bT(v) \forall a, b \in F,\ u, v \in U$.*

Notice that T is a homomorphism of the additive group $(U, +)$ into the additive group $(V, +)$. So you should be able to show that $T(\vec{0}) = \vec{0}$, where the first $\vec{0}$ is the zero vector of U and the second is the zero vector of V.

The zero map: The map $0 : U \to V$ defined by $0(v) = \vec{0}$ for all $v \in U$ is easily seen to be linear.

The identity map: Similarly, the map $I : U \to U$ defined by $I(v) = v$ for all $v \in U$ is linear.

For $f(x) = a_0 + a_1x + a_2x^2 + \cdots a_nx^n \in F[x]$, the formal derivative of $f(x)$ is defined to be $f'(x) = a_1 + 2a_2x + 3a_3x^3 + \cdots + na_nx^{n-1}$.

Differentiation: Define $D : F[x] \to F[x]$ by $D(f) = f'$, where f' is the formal derivative of f. Then D is linear.

The prototypical linear map is given by the following. Let F^n and F^m be the vector spaces of column vectors with n and m entries, respectively. Let $A \in M_{m,n}(F)$. Define $T_A : F^n \to F^m$ by

$$T_A : X \mapsto AX \text{ for all } X \in F^n.$$

The usual properties of matrix algebra force T_A to be linear.

Theorem 4.1.2. *Suppose that U is n-dimensional with basis $\mathcal{B} = (u_1, \ldots, u_n)$, and let $L = (v_1, \ldots v_n)$ be any list of n vectors of V. Then there is a unique linear map $T : U \to V$ such that $T(u_i) = v_i$ for $1 \leq i \leq n$.*

Proof. Let u be any vector of U, so $v = \sum_{i=n}^{n} a_i u_i$ for unique scalars $a_i \in F$. Then the desired T has to be defined by $T(u) = \sum_{i=1}^{n} a_i T(u_i) = \sum_{i=1}^{n} a_i v_i$. This clearly defines T uniquely. The fact that T is linear follows easily from the basic properties of vector spaces. □

Put $\mathcal{L}(U, V) = \{T : U \to V : T \text{ is linear}\}$.

The interesting fact here is that $\mathcal{L}(U, V)$ is again a vector space in its own right. *Vector addition* $S + T$ is defined for $S, T \in \mathcal{L}(U, V)$ by: $(S + T)(u) = S(u) + T(u)$ for all $u \in U$. *Scalar multiplication* aT is defined for $a \in F$ and $T \in \mathcal{L}(U, V)$ by $(aT)(u) = a(T(u))$ for all $u \in U$. You should verify that with this vector addition and scalar multiplication $\mathcal{L}(U, V)$ is a vector space with the zero map being the additive identity.

Now suppose that $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. Then the composition $S \circ T : U \to W$, usually just written ST, defined by $ST(u) = S(T(u))$, is well-defined and is easily shown to be linear. In general the composition product is associative (when a triple product is defined) because this is true of the composition of functions in general. But also we have the distributive properties $(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$ whenever the products are defined.

In general the multiplication of linear maps is not commutative even when both products are defined.

4.2 Kernels and Images

We prefer the following language. If $f : A \to B$ is a function, we say that A is the *domain* of f and B is the *range* of f. But f might not be *onto* B. We define the *image* of f by $\text{Im}(f) = \{b \in B : f(a) = b \text{ for at least one } a \in A\}$.

And we use this language for linear maps also. The *null space* (or *kernel*) of $T \in \mathcal{L}(U, V)$ is defined by $\text{null}(T) = \{u \in U : T(u) = \vec{0}\}$.

In the exercises you are asked to show that the null space and image of a linear map are subspaces of the appropriate spaces.

Lemma 4.2.1. *Let $T \in \mathcal{L}(U, V)$. Then T is injective (i.e., one-to-one) if and only if null$(T) = \{\vec{0}\}$.*

Proof. Since $T(\vec{0}) = \vec{0}$, if T is injective, clearly $\text{null}(T) = \{\vec{0}\}$. Conversely, suppose that $\text{null}(T) = \{\vec{0}\}$. Then suppose that $T(u_1) = T(u_2)$ for $u_1, u_2 \in U$. Then by the linearity of T we have $T(u_1 - u_2) = T(u_1) - T(u_2) = \vec{0}$, so $u_1 - u_2 \in \text{null}(T) = \{\vec{0}\}$. Hence $u_1 = u_2$, implying T is injective. □

The following Theorem *and its method of proof* are extremely useful in many contexts.

Theorem 4.2.2. *If U is finite dimensional and $T \in \mathcal{L}(U, V)$, then Im(T) is finite-dimensional and*

$$dim(U) = dim(null(T)) + dim(Im(T)).$$

Proof. (Pay close attention to the details of this proof. You will want to use them for some of the exercises.)

Start with a basis $(u_1, \ldots, u_k)$ of $\text{null}(T)$. Extend this list to a basis $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ of U. Thus $\dim(\text{null}(T)) = k$ and $\dim(U) = k + r$. To complete a proof of the theorem we need only show that $\dim(\text{Im}(T)) = r$. We will do this by showing that $(T(v_1), \ldots, T(v_r))$ is a basis of $\text{Im}(T)$. Let $u \in U$. Because $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ spans U, there are scalars $a_i, b_i \in F$ such that

$$u = a_1 u_1 + \cdots a_k u_k + b_1 v_1 + \cdots b_r v_r.$$

Remember that $u_1, \ldots, u_k$ are in $\text{null}(T)$ and apply T to both sides of the preceding equation.

$$T(u) = b_1 T(w_1) + \cdots b_r T(w_r).$$

This last equation implies that $(T(w_1), \ldots, T(w_r))$ spans $\text{Im}(T)$, so at least $\text{Im}(T)$ is finite dimensional. To show that $(T(w_1), \ldots, T(w_r))$ is linearly independent, suppose that

$$\sum_{i=1}^{r} c_i T(w_i) = \vec{0} \text{ for some } c_i \in F.$$

It follows easily that $\sum_{i=1}^{r} c_i w_i \in \text{null}(T)$, so $\sum_{i=1}^{r} c_i w_i = \sum_{i=1}^{m} d_i u_i$. Since $(u_1, \ldots, u_m, w_1, \ldots, w_r)$ is linearly independent, we must have that all the c_i's and d_i's are zero. Hence $(T(w_1), \ldots, T(w_r))$ is linearly independent, and hence is a basis for $\text{Im}(T)$. □

Obs. 4.2.3. *There are two other ways to view the equality of the preceding theorem. If $n = dim(U)$, $k = dim(null(T))$ and $r = dim(Im(T))$, then the theorem says $n = k + r$. And if $dim(V) = m$, then clearly $r \leq m$. So $k = n - r \geq n - m$. If $n > m$, then $k > 0$ and T is not injective. If $n < m$, then $r \leq n < m$ says T is not* surjective *(i.e., onto).*

Definition: Often we say that the dimension of the null space of a linear map T is the *nullity of* T.

4.3 Rank and Nullity Applied to Matrices

Let $A \in M_{m,n}(F)$. Recall that the row rank of A (i.e., the dimension of the row space $\text{row}(A)$ of A)equals the column rank (i.e., the dimension of the column space $\text{col}(A)$) of A, and the common value $\text{rank}(A)$ is called the **rank** of A. Let $T_A : F^n \to F^m : X \mapsto AX$ as usual. Then the null space of T_A is the *right null space rnull*(A) of the matrix A, and the image of T_A is the column space of A. So by Theorem 4.2.2 $n = \text{rank}(A) + \dim(\text{rnull}(A))$. Similarly, $m = \text{rank}(A) + \dim(\text{lnull}(A))$. (Clearly $\text{lnull}(A)$ denotes the left null space of A.)

Theorem 4.3.1. *Let A be an $m \times n$ matrix and B be an $n \times p$ matrix over K. Then*

(i) $dim(rnull(AB)) \leq dim(rnull(A)) + dim(rnull(B))$,
and
(ii) $rank(A)$ + $rank(B)$ - $rank(AB) \leq n$.

Proof. Put $V = \text{rnull}(AB) = \{\vec{x} \in K^p : AB\vec{x} = \vec{0} \in K^m\}$. So $\dim(V) = \dim(\text{rnull}(AB)) = p - \text{rank}(AB)$.

Put $W = \text{col}(B) \cap \text{rnull}(A) \subseteq K^n$. Define $T : V \to W : \vec{x} \mapsto B\vec{x}$. As $\text{null}(T)$ is a subspace of $\text{rnull}(B)$, we have the following: $\dim(V) = \dim(\text{rnull}(AB))$; $\dim(W) \leq \dim(\text{rnull}(A))$; and $\dim(\text{null}(T)) \leq \dim(\text{rnull}(B)) = p - \text{rank}(B)$. Then $\dim(V) = \dim(\text{null}(T)) + \dim(\text{Im}(T))$ implies $\dim(\text{rnull}(AB)) = \dim(\text{null}(T)) + \dim(\text{Im}(T)) \leq \dim(\text{rnull}(B)) + \dim(\text{rnull}(A))$, proving (i). This can be rewritten as $p - \text{rank}(AB) \leq (p - \text{rank}(B)) + (n - \text{rank}(A))$, which implies (ii). □

4.4 Projections

Suppose $V = U \oplus W$. Define $P : V \to V$ as follows. For each $v \in V$, write $v = u + w$ with $u \in U$ and $w \in W$. Then u and w are uniquely defined for each $v \in V$. Put $P(v) = u$. It is straightforward to verify the following properties of P.

Obs. 4.4.1. *(i)* $P \in \mathcal{L}(V)$.
(ii) $P^2 = P$.
(iii) $U = Im(P)$.
(iv) $W = null(P)$.
(v) U *and* W *are both* P*-invariant.*

This linear map P is called the *projection onto* U *along* W (or *parallel to* W) and is often denoted by $P = P_{U,W}$. Using this notation we see that

Obs. 4.4.2. $I = P_{U,W} + P_{W,U}$.

As a kind of converse, suppose that $P \in \mathcal{L}(V)$ is *idempotent*, i.e., $P^2 = P$. Put $U = \mathrm{Im}(P)$ and $W = \mathrm{null}(P)$. Then for each $v \in V$ we can write $v = P(v) + (v - P(v)) \in \mathrm{Im}(P) + \mathrm{null}(P)$. Hence $V = U + W$. Now suppose that $u \in U \cap W$. Then on the one hand $u = P(v)$ for some $v \in V$. On the other hand $P(u) = \vec{0}$. Hence $\vec{0} = P(u) = P^2(v) = P(v) = u$, implying $U \cap W = \{\vec{0}\}$. Hence $V = U \oplus W$. It follows readily that $P = P_{U,W}$. Hence we have the following:

Obs. 4.4.3. *The linear map* P *is idempotent if and only if it is the projection onto its image along its null space.*

4.5 Bases and Coordinate Matrices

In this section let V be a finite dimensional vector space over the field F. Since F is a field, i.e., since the multiplication in F is commutative, it turns out that it really does not matter whether V is a *left vector space* over F (i.e., the scalars from F are placed on the left side of the vectors of V) or V is a right vector space. Let the list $\mathcal{B} = (u_1, \ldots, u_n)$ be a basis for V. So if v is an arbitrary vector in V there are unique scalars $c_1, \ldots, c_n$ in F for which

$v = \sum_{i=1}^{n} c_i v_i$. The column vector $[v]_{\mathcal{B}} = (c_1, \ldots, c_n)^T \in F^n$ is then called the *coordinate matrix for v with respect to the basis $\mathcal{B}$*, and we may write

$$v = \sum_{i=1}^{n} c_i v_i = (v_1, \ldots, v_n) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ v_n \end{pmatrix} = \mathcal{B}[v]_{\mathcal{B}}.$$

Perhaps we should discuss this last multiplication a bit.

In the usual theory of matrix manipulation, if we want to multiply two matrices A and B to get a matrix $AB = C$, there are integers n, m and p such that A is $m \times n$, B is $n \times p$, and the product C is $m \times p$. If in general the entry in the ith row and jth column of a matrix A is denoted by A_{ij}, then the (i, j)th entry of $AB = C$ is $(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$. If A is a row or a column the entries are usually indicated by a single subscript. So we might write

$$A = (a_1, \ldots, a_n);\ B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix};\ AB = \sum_{k} a_k b_k.$$

In this context it is usually assumed that the entries of A and B (and hence also of C) come from some ring, probably a commutative ring R, so that in particular this sum $\sum_k a_k b_k$ is a uniquely defined element of R. Moreover, using the usual properties of arithmetic in R it is possible to show directly that matrix multiplication (when defined!) is associative. Also, matrix addition is defined and matrix multiplication (when defined!) distributes over addition, etc. However, it is not always necessary to assume that the entries of A and B come from the same kind of algebraic system. We may just as easily multiply a column $(c_1, \ldots, c_n)^T$ of scalars from the field F by the row $(v_1, \ldots, v_n)$ representing an ordered basis of V over F to obtain

$$v = (v_1, \ldots, v_n) \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \sum_{i=1}^{n} c_i v_i.$$

We may also suppose A is an $n \times n$ matrix over F and write

$$(u_1, u_2, \ldots, u_n) = (v_1, v_2, \ldots, v_n)A, \text{ so } u_j = \sum_{i=1}^{n} A_{ij} v_i.$$

Then it follows readily that $(u_1, \ldots, u_n)$ is an ordered basis of V if and only if the matrix A is invertible, in which case it is also true that

$$(v_1, v_2, \ldots, v_n) = (u_1, u_2, \ldots, u_n)A^{-1}, \text{ so } v_j = \sum_{i=1}^{n} A^{-1}_{ij} u_i.$$

4.6 Matrices as Linear Transformations

Let $\mathcal{B}_1 = (u_1, \ldots, u_n)$ be an ordered basis for the vector space U over the field F, and let $\mathcal{B}_2 = (v_1, \ldots, v_m)$ be an ordered basis for the vector space V over F. Let A be an $m \times n$ matrix over F. Define $T_A : U \to V$ by $[T_A(u)]_{\mathcal{B}_2} = A \cdot [u]_{\mathcal{B}_1}$ for all $u \in U$. It is quite straightforward to show that $T_A \in \mathcal{L}(U, V)$. It is also clear (by letting $u = u_j$), that the j^{th} column of A is $[T(u_j)]_{\mathcal{B}_2}$. Conversely, if $T \in \mathcal{L}(U, V)$, and if we define the matrix A to be the matrix with j^{th} column equal to $[T(u_j)]_{\mathcal{B}_2}$, then $[T(u)]_{\mathcal{B}_2} = A \cdot [u]_{\mathcal{B}_1}$ for all $u \in U$. In this case we say that *A is the matrix that represents T with respect to the pair $(\mathcal{B}_1, \mathcal{B}_2)$ of ordered bases of U and V*, respectively, and we write $A = [T]_{\mathcal{B}_2, \mathcal{B}_1}$.

Note that a coordinate matrix of a vector with respect to a basis has a subscript that is a single basis, whereas the matrix representing a linear map T has a subscript which is a pair of bases, with the basis of the range space listed first and that of the domain space listed second. Soon we shall see why this order is the convenient one. When $U = V$ and $\mathcal{B}_1 = \mathcal{B}_2$ it is sometimes the case that we write $[T]_{\mathcal{B}_1}$ in place of $[T]_{\mathcal{B}_1, \mathcal{B}_1}$. And we usually write $\mathcal{L}(V)$ in place of $\mathcal{L}(V, V)$, and $T \in \mathcal{L}(V)$ is called a *linear operator on V*.

In these notes, however, even when there is only one basis of V being used and T is a linear operator on V, we sometimes indicate the matrix that represents T with a subscript that is a *pair* of bases, instead of just one basis, because there are times when we want to think of T as a member of a vector space so that it has a coordinate matrix with respect to some basis of that vector space. Our convention makes it easy to recognize when the matrix represents T with respect to a basis as a linear map and when it represents T as a vector itself which is a linear combination of the elements of some basis.

We give an example of this.

Example 4.6.1. *Let $V = M_{2,3}(F)$. Put $\mathcal{B} = (v_1, \ldots, v_6)$ where the v_i are defined as follows:*

$$v_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; v_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}; v_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix};$$

$$v_4 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}; v_5 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}; v_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is clear that $\mathcal{B}$ is a basis for $M_{2,3}(F)$, and if $A \in M_{2,3}(F)$, then the coordinate matrix $[A]_{\mathcal{B}}$ of A with respect to the basis $\mathcal{B}$ is

$$[A]_{\mathcal{B}} = (A_{11}, A_{12}, A_{13}, A_{21}, A_{22}, A_{23})^T.$$

Given such a matrix A, define a linear map $T_A : F^3 \to F^2$ by $T_A(u) = Au$ for all $u \in F^3$. Let $\mathcal{S}_1 = (e_1, e_2, e_3)$ be the standard ordered basis of F^3, and let $\mathcal{S}_2 = (h_1, h_2)$ be the standard ordered basis of F^2. So, for example, $h_2 = (0, 1)$. You should check that

$$[T_A]_{\mathcal{S}_2, \mathcal{S}_1} = A.$$

For $1 \leq i \leq 3$; $1 \leq j \leq 2$, let $f_{ij} \in \mathcal{L}(F^3, F^2)$ be defined by $f_{ij}(e_k) = \delta_{ik} h_j$. Let $\mathcal{B}_3 = (f_{11}, f_{21}, f_{31}, f_{12}, f_{22}, f_{32})$. We want to figure out what is the coordinate matrix $[T_A]_{\mathcal{B}_3}$.

We claim that $[T_A]_{\mathcal{B}_3} = (a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23})^T$. Because of the order in which we listed the basis vectors f_{ij}, this is equivalent to saying that $T_A = \sum_{i,j} a_{ij} f_{ji}$. If we evaluate this sum at (e_k) we get

$$\sum_{i,j} a_{ij} f_{ji}(e_k) = \sum_i a_{ik} f_{ki}(e_k) = \sum_i a_{ik} h_i = T_A(e_k).$$

This establishes our claim.

Recall that $\mathcal{B}_1 = (u_1, \ldots, u_n)$ is an ordered basis for U over the field F, and that $\mathcal{B}_2 = (v_1, \ldots, v_m)$ is an ordered basis for V over F. Now suppose that W has an ordered basis $\mathcal{B}_3 = (w_1, \ldots, w_p)$.

Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$, so that $T \circ S \in \mathcal{L}(U, W)$, where $T \circ S$ means do S first. Then we have

$$[(T \circ S)(u)]_{\mathcal{B}_3} = [T(S(u))]_{\mathcal{B}_3} = [T]_{\mathcal{B}_3, \mathcal{B}_2} [S(u)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_3, \mathcal{B}_2} [S]_{\mathcal{B}_2, \mathcal{B}_1} [u]_{\mathcal{B}_1} =$$

$$= [T \circ S]_{\mathcal{B}_3,\mathcal{B}_1}[u]_{\mathcal{B}_1} \text{ for all } u \in U.$$

This implies that

$$[T \circ S]_{\mathcal{B}_3,\mathcal{B}_1} = [T]_{\mathcal{B}_3,\mathcal{B}_2} \cdot [S]_{\mathcal{B}_2,\mathcal{B}_1}.$$

This is the equation that suggests that the subscript on the matrix representing a linear map should have the basis for the range space listed first.

Recall that $\mathcal{L}(U,V)$ is naturally a vector space over F with the usual addition of linear maps and scalar multiplication of linear maps. Moreover, for $a, b \in F$ and $S, T \in \mathcal{L}(U,V)$, it follows easily that

$$[aS + bT]_{\mathcal{B}_2,\mathcal{B}_1} = a[S]_{\mathcal{B}_2,\mathcal{B}_1} + b[T]_{\mathcal{B}_2,\mathcal{B}_1}.$$

We leave the proof of this fact as a straightforward exercise. It then follows that if $U = V$ and $\mathcal{B}_1 = \mathcal{B}_2$, the correspondence $T \mapsto [T]_{\mathcal{B}_1,\mathcal{B}_1} = [T]_{\mathcal{B}_1}$ is an algebra isomorphism. This includes consequences such as $[T^{-1}]_{\mathcal{B}} = ([T]_{\mathcal{B}})^{-1}$ when T happens to be invertible. Proving these facts is a worthwhile exercise!

Let $f_{ij} \in \mathcal{L}(U,V)$ be defined by

$$f_{ij}(u_k) = \delta_{ik} v_j, \quad 1 \le i,k \le n;\ 1 \le j \le m.$$

So f_{ij} maps u_i to v_j and maps u_k to the zero vector for $k \neq i$. This completely determines f_{ij} as a linear map from U to V.

Theorem 4.6.2. *The set $\mathcal{B}^* = \{f_{ij} : 1 \le i \le m; 1 \le j \le n\}$ is a basis for $\mathcal{L}(U,V)$ as a vector space over F.*

Note: We could turn $\mathcal{B}^*$ into a list, but we don't need to.

Proof. We start by showing that $\mathcal{B}^*$ is linearly independent. Suppose that $\sum_{ij} c_{ij} f_{ij} = 0$, so that $\vec{0} = \sum_{ij} c_{ij} f_{ij}(u_k) = \sum_j c_{kj} v_j$ for each k. Since $(v_1, \ldots, v_m)$ is linearly independent, $c_{k1} = c_{k2} = \cdots = c_{km} = 0$, and this holds for each k, so the set of f_{ij} must be linearly independent. We now show that it spans $\mathcal{L}(U,V)$. For suppose that $S \in \mathcal{L}(U,V)$ and that $[S]_{\mathcal{B}_2,\mathcal{B}_1} = C = (c_{ij})$, i.e., $S(u_j) = \sum_{i=1}^n c_{ij} v_i$. Put $T = \sum_{i,k} c_{ik} f_{ki}$. Then $T(u_j) = \sum c_{ik} f_{ki}(u_j) = \sum_i c_{ij} v_i$, implying that $S = T$ since they agree on a basis. □

Corollary 4.6.3. *If $dim(U) = n$ and $dim(V) = m$, then $dim\mathcal{L}(U,V) = mn$.*

4.7 Change of Basis

We want to investigate what happens to coordinate matrices and to matrices representing linear operators when the ordered basis is changed. For the sake of simplicity we shall consider this question only for linear operators on a space V, so that we can use a single ordered basis. So in this section we write matrices representing linear operators with a subscript which is a single basis.

Let F be any field and let V be a finite dimensional vector space over F, say $dim(V) = n$. Let $\mathcal{B}_1 = (u_1, \ldots, u_n)$ and $\mathcal{B}_2 = (v_1, \ldots, v_n)$ be two (ordered) bases of V. So for $v \in V$, and for $i = 1$, say that $[v]_{\mathcal{B}_1} = (c_1, c_2, \ldots, c_n)^T$, i.e., $v = \sum_{i=1}^{n} c_i u_i$. We often write this equality in the form

$$v = (u_1, \ldots, u_n)[v]_{\mathcal{B}_1} = \mathcal{B}_1[v]_{\mathcal{B}_1}.$$

Similarly, $v = \mathcal{B}_2[v]_{\mathcal{B}_2}$.

Since $\mathcal{B}_1$ and $\mathcal{B}_2$ are both bases for V, there is an invertible matrix Q such that

$$\mathcal{B}_1 = \mathcal{B}_2 Q \text{ and } \mathcal{B}_2 = \mathcal{B}_1 Q^{-1}.$$

The first equality indicates that

$$u_j = \sum_{i=1}^{n} Q_{ij} v_i. \tag{4.1}$$

This equation says that the jth column of Q is the coordinate matrix of u_j with respect to $\mathcal{B}_2$. Similarly, the jth column of Q^{-1} is the coordinate matrix of v_j with respect to $\mathcal{B}_1$.

For every $v \in V$ we now have

$$v = \mathcal{B}_1[v]_{\mathcal{B}_1} = (\mathcal{B}_2 Q)[v]_{\mathcal{B}_1} = \mathcal{B}_2[v]_{\mathcal{B}_2}.$$

It follows that

$$Q[v]_{\mathcal{B}_1} = [v]_{\mathcal{B}_2}. \tag{4.2}$$

Now let $T \in \mathcal{L}(V)$. Recall that the matrix $[T]_{\mathcal{B}}$ that represents T with respect to the basis $\mathcal{B}$ is the unique matrix for which

$$[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}} \text{ for all } v \in V.$$

Theorem 4.7.1. *Let $\mathcal{B}_1 = \mathcal{B}_2 Q$ as above. Then $[T]_{\mathcal{B}_2} = Q[T]_{\mathcal{B}_1} Q^{-1}$.*

Proof. In Eq. 4.2 replace v with $T(v)$ to get

$$Q[T(v)]_{\mathcal{B}_1} = [T(v)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_2}[v]_{\mathcal{B}_2} = [T]_{\mathcal{B}_2} Q[v]_{\mathcal{B}_1}$$

for all $v \in V$. It follows that

$$Q[T]_{\mathcal{B}_1}[v]_{\mathcal{B}_1} = [T]_{\mathcal{B}_2} Q[v]_{\mathcal{B}_1}$$

for all $v \in V$, implying

$$Q[T]_{\mathcal{B}_1} = [T]_{\mathcal{B}_2} Q, \tag{4.3}$$

which is equivalent to the statement of the theorem. □

A Specific Setting

Now let $V = F^n$, whose elements we think of as being column vectors. Let A be an $n \times n$ matrix over F and define $T_A \in \mathcal{L}(F^n)$ by

$$T_A(v) = Av, \text{ for all } v \in F^n.$$

Let $\mathcal{S} = (e_1, \ldots, e_n)$ be the standard ordered basis for F^n, i.e., e_j is the column vector in F^n whose jth entry is 1 and all other entries are equal to 0. It is clear that we can identify each vector $v \in F^n$ with $[v]_{\mathcal{S}}$. Moreover, the jth column of A is $Ae_j = [Ae_j]_{\mathcal{S}} = [T_A e_j]_{\mathcal{S}} = [T_A]_{\mathcal{S}}[e_j]_{\mathcal{S}} = [T_A]_{\mathcal{S}} e_j =$ the jth column of $[T_A]_{\mathcal{S}}$, which implies that

$$A = [T_A]_{\mathcal{S}}. \tag{4.4}$$

Theorem 4.7.2. *Let $\mathcal{S} = (e_1, \ldots, e_n)$ be the standard ordered basis of F^n. Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a second ordered basis. Let P be the matrix whose jth column is $v_j = [v_j]_{\mathcal{S}}$. Let A be an $n \times n$ matrix over F and define $T_A : F^n \to F^n$ by $T_A(v) = Av$. So $[T_A]_{\mathcal{S}} = A$. Then $[T_A]_{\mathcal{B}} = P^{-1}AP$.*

Proof. Since $v_j = [v_j]_{\mathcal{S}} = \mathcal{S}[v_j]_{\mathcal{S}}$, it follows that

$$(v_1, \ldots, v_n) = \mathcal{S}([v_1]_{\mathcal{S}}, \ldots, [v_n]_{\mathcal{S}}) = \mathcal{S}P,$$

i.e., $\mathcal{B} = \mathcal{S}P$, which is equivalent to $\mathcal{S} = \mathcal{B}P^{-1}$. So if $\mathcal{S}$ plays the role of $\mathcal{B}_1$ above, and $\mathcal{B}$ plays the role of $\mathcal{B}_2$, and P^{-1} plays the role of Q, we have

$$[v]_{\mathcal{B}} = P^{-1}[v]_{\mathcal{S}}, \tag{4.5}$$

$$[T_A]_{\mathcal{B}} = P^{-1}[T_A]_{\mathcal{S}}P = P^{-1}AP. \tag{4.6}$$

□

Definition Two $n \times n$ matrices A and B over F are said to be *similar* (written $A \sim B$) if and only if there is an invertible $n \times n$ matrix P such that $B = P^{-1}AP$. You should prove that "similarity" is an equivalence relation on $M_n(F)$.

Corollary 4.7.3. *If $A,\ B \in M_n(F)$, and if V is an n-dimensional vector space over F, then A and B are similar if and only if there are bases $\mathcal{B}_1$ and $\mathcal{B}_2$ of V and $T \in \mathcal{L}(V)$ such that $A = [T]_{\mathcal{B}_1}$ and $B = [T]_{\mathcal{B}_2}$.*

The Dual Space[1]

We now specialize to the case where $V = F$ is viewed as a vector space over F. Here $\mathcal{L}(U, F)$ is denoted U^* and is called the *dual space* of U. An element of $\mathcal{L}(U, F)$ is called a *linear functional*. Write $\mathcal{B} = (u_1, \ldots, u_n)$ for the fixed ordered basis of U. Then $1 \in F$ is a basis of F over F, so we write $\bar{1} = (1)$ and $m = 1$, and there is a basis $\mathcal{B}^*$ of U^* defined by $\mathcal{B}^* = (f_1, \ldots, f_n)$, where $f_i(u_j) = \delta_{ij} \in F$. This basis $\mathcal{B}^*$ is called the *basis dual to* $\mathcal{B}$. If $f \in U^*$ satisfies $f(u_j) = c_j$ for $1 \leq j \leq n$, then

$$[f]_{\bar{1},\mathcal{B}_1} = [c_1, \ldots, c_n] = [f(u_1), \ldots, f(u_n)].$$

As above in the more general case, if $g = \sum_i c_i f_i$, then $g(u_j) = \sum_i c_i f_i(u_j) = c_j$, so $f = g$, and $[f]_{\mathcal{B}^*} = [c_1, \ldots, c_n]^T = ([f]_{\bar{1},\mathcal{B}_1})^T$. We restate this in general:

$$\text{For } f \in U^*, [f]_{\mathcal{B}^*} = ([f]_{\bar{1},\mathcal{B}_1})^T.$$

Now we suppose that $T \in GL(U)$, i.e., T is an invertible element of $\mathcal{L}(U, U)$. We define a map $\hat{T} : U^* \to U^*$ by

$$\hat{T}(f) = f \circ T^{-1}, \quad \text{for all } f \in U^*.$$

We want to determine the matrix $[\hat{T}]_{\mathcal{B}^*,\mathcal{B}^*}$. We know that the j^{th} column of this matrix is $[\hat{T}(f_j)]_{\mathcal{B}^*} = [f_j \circ T^{-1}]_{\mathcal{B}^*} = \left([f_j \circ T^{-1}]_{\bar{1},\mathcal{B}}\right)^T = \left([f_j]_{\bar{1},\mathcal{B}} \cdot [T^{-1}]_{\mathcal{B},\mathcal{B}}\right)^T =$

$$\left((0, \cdots, 1_j, \cdots, 0)\,([T]_{\mathcal{B},\mathcal{B}})^{-1}\right)^T = ([T]_{\mathcal{B},\mathcal{B}})^{-T} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1_j \\ \vdots \\ 0 \end{pmatrix} = j^{th} \text{ column of } ([T]_{\mathcal{B},\mathcal{B}})^{-T}.$$

[1] Note that this subsection on the dual space may be omitted.

This says that

$$[\hat{T}]_{\mathcal{B}^*,\mathcal{B}^*} = ([T]_{\mathcal{B},\mathcal{B}})^{-T}.$$

4.8 Exercises

1. Let $T \in \mathcal{L}(U,V)$. Then $\text{null}(T)$ is a subspace of U and $\text{Im}(T)$ is a subspace of V.

2. Suppose that V and W are finite-dimensional and that U is a subspace of V. Prove that there exists a $T \in \mathcal{L}(V,W)$ such that $\text{null}(T) = U$ if and only if $\dim(U) \geq \dim(V) - \dim(W)$.

3. Let $T \in \mathcal{L}(V)$. Put $R = \text{Im}(T)$ and $N = \text{null}(T)$. Note that both R and N are T-invariant. Show that R has a complementary T-invariant subspace W (i.e., $V = R \oplus W$ and $T(W) \subseteq W$) if and only if $R \cap N = \{\vec{0}\}$, in which case N is the unique T-invariant subspace complementary to R.

4. State and prove Theorem 4.3.1 for linear maps (instead of for matrices).

5. Prove Corollary 4.6.3

6. If $T =\in \mathcal{L}(U,V)$, we know that as a function from U to V, T has an inverse if and only if it is *bijective* (i.e., one-to-one and onto). Show that when T is invertible as a function, then its inverse is in $\mathcal{L}(V,U)$.

7. If $T \in \mathcal{L}(U,V)$ is invertible, and if $\mathcal{B}_1$ is a basis for U and $\mathcal{B}_2$ is a basis for V, then $([T]_{\mathcal{B}_2,\mathcal{B}_1})^{-1} = [T^{-1}]_{\mathcal{B}_1,\mathcal{B}_2}$.

8. Two vector spaces U and V are said to be *isomorphic* provided there is an invertible $T \in \mathcal{L}(U,V)$. Show that if U and V are finite-dimensional vector spaces over F, then U and V are isomorphic if and only if $\dim(U) = \dim(V)$.

9. Let $A \in M_{m,n}(F)$ and $b \in M_{m,1}(F) = F^m$. Consider the matrix equation $A\vec{x} = \vec{b}$ as a system of m linear equations in n unknowns $x_1, \ldots, x_n$. Interpret Obs. 4.2.3 for this system of linear equations.

10. Let $V = M_{2,3}(F)$. Suppose that $B \in M_{2,2}(F)$, say that $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ and also that $C \in M_{3,3}(F)$ with (i,j)-entry equal to c_{ij}. Write $\mathcal{L}(V)$ for $\mathcal{L}(V,V)$ and define $T_{B,C} \in \mathcal{L}(V)$ by

$$T_{B,C} : A \mapsto BAC \text{ for all } A \in M_{2,3}(F).$$

First verify that $T_{B,C} \in \mathcal{L}(V)$. Then construct a basis $\mathcal{B}_1$ of $\mathcal{L}(V)$ and determine the coordinate matrix $[T_{B,C}]_{\mathcal{B}_1}$. Let $\mathcal{B}_3$ be the basis of $\mathcal{L}(F^3, F^2)$ used in Example 4.6.1. Then compute the matrix $[T_{B,C}]_{\mathcal{B}_3,\mathcal{B}_3}$ that represents $T_{B,C}$.

11. Let $\mathcal{B}_1$, $\mathcal{B}_2$ be bases for U and V, respectively, with $\dim(U) = n$ and $\dim(V) = m$. Show that the map

$$\mathcal{M} : \mathcal{L}(U,V) \to M_{m,n}(F) : T \mapsto [T]_{\mathcal{B}_2,\mathcal{B}_1}$$

is an invertible linear map.

12. Suppose that V is finite-dimensional and that $T \in \mathcal{L}(V)$. Show that the following are equivalent:

(i) T is invertible.

(ii) T is injective.

(iii) T is surjective.

13. Suppose that V is finite dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST is invertible if and only if both S and T are invertible.

14. Suppose that V is finite dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

15. Suppose that W is finite dimensional and $T \in \mathcal{L}(V,W)$. Prove that T is injective if and only if there exists an $S \in \mathcal{L}(W,V)$ such that ST is the identity map on V.

16. Suppose that V is finite dimensional and $T \in \mathcal{L}(V,W)$. Prove that T is surjective if and only if there exists an $S \in \mathcal{L}(W,V)$ such that TS is the identity map on W.

Chapter 5

Polynomials

5.1 Algebras

It is often the case that basic facts about polynomials are taken for granted as being well-known and the subject is never developed in a formal manner. In this chapter, which we usually assign as independent reading, we wish to give the student a somewhat formal introduction to the algebra of polynomials over a field. It is then natural to generalize to polynomials with coefficients from some more general algebraic structure, such as a commutative ring. The title of the course for which this book is intended includes the words "linear algebra," so we feel some obligation to define what a linear algebra is.

Definition Let F be a field. A *linear algebra over the field* F is a vector space $\mathcal{A}$ over F with an additional operation called *multiplication of vectors* which associates with each pair of vectors $u, v \in \mathcal{A}$ a vector uv in $\mathcal{A}$ called the *product* of u and v in such a way that

(a) multiplication is associative: $u(vw) = (uv)w$ for all $u, v, w \in \mathcal{A}$;

(b) multiplication distributes over addition: $u(v + w) = (uv) + (uw)$ and $(u + v)w = (uw) + (vw)$, for all $u, v, w \in \mathcal{A}$;

(c) for each scalar $c \in F$, $c(uv) = (cu)v = u(cv)$ for all $u, v \in \mathcal{A}$.

If there is an element $1 \in \mathcal{A}$ such that $1u = u1 = u$ for each $u \in \mathcal{A}$, we call $\mathcal{A}$ a *linear algebra with identity over* F, and call 1 the *identity* of $\mathcal{A}$. The algebra $\mathcal{A}$ is called *commutative* provided $uv = vu$ for all $u, v \in \mathcal{A}$.

Example 5.1.1. *The set of $n \times n$ matrices over a field, with the usual operations, is a linear algebra with identity; in particular the field itself is an algebra with identity. This algebra is not commutative if $n \geq 2$. Of course,*

the field itself is commutative.

Example 5.1.2. *The space of all linear operators on a vector space, with composition as the product, is a linear algebra with identity. It is commutative if and only if the space is one-dimensional.*

Now we turn our attention to the construction of an algebra which is quite different from the two just given. Let F be a field and let S be the set of all nonnegative integers. We have seen that the set of all functions from S into F is a vector space which we now denote by F^∞. The vectors in F^∞ are just infinite sequences (i.e., lists) $f = (f_0, f_1, f_2, \ldots)$ of scalars $f_i \in F$. If $g = (g_0, g_1, g_2, \ldots)$ and $a, b \in F$, then $af + bg$ is the infinite list given by

$$af + bg = (af_0 + bg_0, af_1 + bg_1, \ldots) \tag{5.1}$$

We define a product in F^∞ by associating with each pair (f, g) of vectors in F^∞ the vector fg which is given by

$$(fg)_n = \sum_{i=1}^{n} f_i g_{n-i}, \quad n = 0, 1, 2, \ldots \tag{5.2}$$

Since multiplication in F is commutative, it is easy to show that multiplication in F^∞ is also commutative. In fact, it is a relatively routine task to show that F^∞ is now a linear algebra with identity over F. Of course the vector $(1, 0, 0, \ldots)$ is the identity, and the vector $x = (0, 1, 0, 0, \ldots)$ plays a distinguished role. Throughout this chapter x will continue to denote this particular vector (and will never be an element of the field F). The product of x with itself n times will be denoted by x^n, and by convention $x^0 = 1$. Then

$$x^2 = (0, 0, 1, 0, \ldots), \quad x^3 = (0, 0, 0, 1, 0, \ldots), \text{ etc.}$$

Obs. 5.1.3. *The list* $(1, x, x^2, \ldots)$ *is both independent and infinite. Thus the algebra* F^∞ *is not finite dimensional.*

The algebra F^∞ is sometimes called the *algebra of formal power series over* F. The element $f = (f_0, f_1, f_2, \ldots)$ is frequently written as

$$f = \sum_{n=0}^{\infty} f_n x^n. \tag{5.3}$$

This notation is very convenient, but it must be remembered that it is purely formal. In algebra there is no such thing as an 'infinite sum,' and the power series notation is not intended to suggest anything about convergence.

5.2 The Algebra of Polynomials

Definition Let $F[x]$ be the subspace of F^∞ spanned by the vectors $1, x, x^2, \ldots,$. An element of $F[x]$ is called a *polynomial over* F.

Since $F[x]$ consists of all (finite) linear combinations of x and its powers, a non-zero vector f in F^∞ is a polynomial if and only if there is an integer $n \geq 0$ such that $f_n \neq 0$ and such that $f_k = 0$ for all integers $k > n$. This integer (when it exists) is called the *degree* of f and is denoted by $\deg(f)$.The zero polynomial is said to have degree $-\infty$. So if $f \in F[x]$ has degree n, it may be written in the form

$$f = f_0x^0 + f_1x + f_2x^2 + \cdots + f_nx^n, \quad f_n \neq 0.$$

Usually f_0x^0 is simply written f_0 and called a *scalar polynomial.* A non-zero polynomial f of degree n such that $f_n = 1$ is called a *monic* polynomial. The verification of the various parts of the next result guaranteeing that $\mathcal{A}$ is an algebra is routine and is left to the reader.

Theorem 5.2.1. *Let f and g be non-zero polynomials over F. Then*

(i) fg is a non-zero polynomial;

(ii) $deg(fg) = deg(f) + deg(g)$;

(iii) fg is a monic polynomial if both f and g are monic;

(iv) fg is a scalar polynomial if and only if both f and g are scalar polynomials;

(v) $deg(f + g) \leq max\{deg(f), deg(g)\}$.

Corollary 5.2.2. *The set $F[x]$ of all polynomials over a given field F with the addition and multiplication given above is a commutative linear algebra with identity over F.*

Corollary 5.2.3. *Suppose f, g, and h are polynomials over F such that $f \neq 0$ and $fg = fh$. Then $g = h$.*

Proof. Since $fg = fh$, also $f(g - h) = 0$. Since $f \neq 0$, it follows from (i) above that $g - h = 0$. □

Let $f = \sum_{i=0}^{m} f_ix^i$ and $g = \sum_{j=0}^{n} g_jx^j$, and interpret $f_k = 0$, $g_t = 0$, if $k > m$, $t > n$, respectively. Then

$$fg = \sum_{i,j} f_ig_jx^{i+j} = \sum_{i=0}^{m+n} \left(\sum_{j=0}^{i} f_jg_{i-j} \right) x^i,$$

where the first sum is extended over all integer pairs (i, j) with $0 \leq i \leq m$ and $0 \leq j \leq n$.

Definition Let $\mathcal{A}$ be a linear algebra with identity over the field F. We denote the identity of $\mathcal{A}$ by 1 and make the convention that $u^0 = 1$ for each $u \in \mathcal{A}$. Then to each polynomial $f = \sum_{i=1}^{n} f_i x^i$ over F and $u \in \mathcal{A}$, we associate an element $f(u)$ in $\mathcal{A}$ by the rule

$$f(u) = \sum_{i=0}^{n} f_i u^i.$$

Example 5.2.4. *Let $\mathcal{C}$ be the field of complex numbers and let $f = x^2 + 2$.*

(a) If $\mathcal{A} = \mathcal{C}$ and $z \in \mathcal{C}$, $f(z) = z^2 + 2$, in particular $f(3) = 11$ and

$$f\left(\frac{1+i}{1-i}\right) = 1.$$

(b) If $\mathcal{A}$ is the algebra of all 2×2 matrices over $\mathcal{C}$ and if

$$B = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix},$$

then

$$f(B) = 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}^2 = \begin{pmatrix} 3 & 0 \\ -3 & 6 \end{pmatrix}.$$

(c) If $\mathcal{A}$ is the algebra of all liinear operators on $\mathcal{C}^3$ and T is the element of $\mathcal{A}$ given by

$$T(c_1, c_2, c_3) = (i\sqrt{2}c_1, c_2, i\sqrt{2}c_3),$$

then $f(T)$ is the linear operator on $\mathcal{C}^3$ defined by

$$f(T)(c_1, c_2, c_3) = (0, 3c_2, 0).$$

(d) If $\mathcal{A}$ is the algebra of all polynomials over $\mathcal{C}$ and $t = x^4 + 3i$, then $f(g)$ is the polynomial in $\mathcal{A}$ given by

$$f(g) = -7 + 6ix^4 + x^8.$$

Theorem 5.2.5. *Let F be a field, $\mathcal{A}$ a linear algebra with identity over F, $f, g \in F[x]$, $u \in \mathcal{A}$ and $c \in F$. Then:*
(i) $(cf + g)(u) = cf(u) + g(u)$;
(ii) $(fg)(u) = f(u)g(u) = (gf)(u)$.

Proof. We leave (i) as an exercise. So for (ii), suppose

$$f = \sum_{i=0}^{m} f_i x^i \text{ and } g = \sum_{j=0}^{n} g_j x^j.$$

Recall that $fg = \sum_{i,j} f_i g_j x^{i+j}$. So using (i) we obtain

$$(fg)(u) = \sum_{i,j} f_i g_j u^{i+j} = \left(\sum_{i=0}^{m} f_i u^i\right)\left(\sum_{j=0}^{n} g_j u^j\right) = f(u)g(u).$$

□

Fix $u \in \mathcal{A}$ and define $E_u : F[x] \to \mathcal{A}$ by

$$E_u(f) = f(u). \tag{5.4}$$

Using Theorem 5.2.5 it is now easy to see that the map $E_u : F[x] \to \mathcal{A}$ is an algebra homomorphism, i.e., it preserves addition and multiplication. There is a special case of this that is so important for us that we state it as a separate corollary.

Corollary 5.2.6. *If $\mathcal{A} = \mathcal{L}(V)$ and $T \in \mathcal{A}$, and if $f, g \in F[x]$,then* $(f \cdot g)(T) = f(T) \circ g(T)$.

5.3 Lagrange Interpolation

Throughout this section F is a fixed field and $t_0, t_1, \ldots, t_n$ are $n+1$ *distinct* elements of F. Put $V = \{f \in F[x] : \deg(f) \leq n\}$, and define $E_i : V \to F$ by $E_i(f) = f(t_i)$, $0 \leq i \leq n$.

By Theorem 5.2.5 each E_i is a linear functional on V. Moreover, we show that $\mathcal{B}^* = (E_0, E_1, \ldots, E_n)$ is the basis of V^* dual to a particular basis of V.

Put

$$p_i = \prod_{j \neq i} \left(\frac{x - t_j}{t_i - t_j}\right).$$

Then each p_i has degree n, so belongs to V, and

$$E_j(p_i) = p_i(t_j) = \delta_{ij}. \tag{5.5}$$

It will turn out that $\mathcal{B} = (p_0, \ldots, p_n)$ is a basis for V, and then Eq. 5.5 expresses what we mean by saying that $\mathcal{B}^*$ is the basis dual to $\mathcal{B}$.

If $f = \sum_{i=0}^{n} c_i p_i$, then for each j

$$f(t_j) = \sum_i c_i p_i(t_j) = c_j. \tag{5.6}$$

So if f is the zero polynomial, each c_j must equal 0, implying that the list $(p_0, \ldots, p_n)$ is linearly independent in V. Since $(1, x, x^2, \ldots, x^n)$ is a basis for V, clearly $\dim(V) = n+1$. Hence $\mathcal{B} = (p_0, \ldots, p_n)$ must be a basis for V. It then follows from Eq. 5.6 that for each $f \in V$, we have

$$f = \sum_{i=0}^{n} f(t_i) p_i. \tag{5.7}$$

The expression in Eq. 5.7 is known as **Lagrange's Interpolation Formula**. Setting $f = x^j$ in Eq. 5.7 we obtain

$$x^j = \sum_{i=0}^{n} (t_i)^j p_i \tag{5.8}$$

Definition Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two linear algebras over F. They are said to be *isomorphic* provided there is a one-to-one mapping $u \mapsto u'$ of $\mathcal{A}_1$ onto $\mathcal{A}_2$ such that

(a) $(cu + dv)' = cu' + dv'$

and

(b) $(uv)' = u'v'$

for all $u,\ v \in \mathcal{A}_1$ and all scalars $c,\ d \in F$. The mapping $u \mapsto u'$ is called an *isomorphism* of $\mathcal{A}_1$ onto $\mathcal{A}_2$. An isomorphism of $\mathcal{A}_1$ onto $\mathcal{A}_2$ is thus a vector space isomorphism of $\mathcal{A}_1$ onto $\mathcal{A}_2$ which has the additional property of preserving products.

Example 5.3.1. *Let V be an n-dimensional vector space over the field F. As we have seen earlier, each ordered basis $\mathcal{B}$ of V determines an* isomorphism *$T \mapsto [T]_{\mathcal{B}}$ of the algebra of linear operators on V onto the algebra of $n \times n$*

matrices over F. Suppose now that S is a fixed linear operator on V and that we are given a polynomial

$$f = \sum_{i=0}^{n} c_i x^i$$

with coefficients $c_i \in F$. Then

$$f(S) = \sum_{i=0}^{n} c_i S^i.$$

Since $T \mapsto [T]_{\mathcal{B}}$ is a linear mapping,

$$[f(S)]_{\mathcal{B}} = \sum_{i=0}^{n} c_i [S^i]_{\mathcal{B}}.$$

From the additional fact that

$$[T_1 T_2]_{\mathcal{B}} = [T_1]_{\mathcal{B}} [T_2]_{\mathcal{B}}$$

for all T_1, $T_2 \in \mathcal{L}(V)$, it follows that

$$[S^i]_{\mathcal{B}} = \left([S]_{\mathcal{B}}\right)^i, \quad 2 \leq i \leq n.$$

As this relation is also valid for $i = 0$ and 1, we obtain the result that

Obs. 5.3.2.

$$[f(S)]_{\mathcal{B}} = f\left([S]_{\mathcal{B}}\right).$$

In other words, if $S \in \mathcal{L}(V)$, the matrix of a polynomial in S, with respect to a given basis, is the same polynomial in the matrix of S.

5.4 Polynomial Ideals

In this section we are concerned primarily with the fact that $F[x]$ is a principal ideal domain.

Lemma 5.4.1. *Suppose f and d are non-zero polynomials in $F[x]$ such that $deg(d) \leq deg(f)$. Then there exists a poynomial $g \in F[x]$ for which*

$$deg(f - dg) < deg(f).$$

Note: This includes the possibility that $f = dg$ so $\deg(f - dg) = -\infty$.

Proof. Suppose

$$f = a_m x^m + \sum_{i=0}^{m-1} a_i x^i, \quad a_m \neq 0$$

and that

$$d = b_n x^n + \sum_{i=0}^{n-1} b_i x^i, \quad b_n \neq 0.$$

Then $m \geq n$ and

$$f - \left(\frac{a_m}{b_n}\right) x^{m-n} d = 0 \text{ or } \quad \deg\left[f - \left(\frac{a_m}{b_n}\right) x^{m-n} d\right] < \deg(f).$$

Thus we may take $g = \left(\frac{a_m}{b_n}\right) x^{m-n}$. □

This lemma is useful in showing that the usual algorithm for "long division" of polynomials works over any field.

Theorem 5.4.2. *If $f, d \in F[x]$ and $d \neq 0$, then there are unique polynomials $q, r \in F[x]$ such that*
(i) $f = dq + r$;
(ii) $deg(r) < deg(d)$.

Proof. If $\deg(f) < \deg(d)$ we may take $q = 0$ and $r = f$. In case $f \neq 0$ and $\deg(f) \geq \deg(d)$, the preceding lemma shows that we may choose a polynomial $g \in F[x]$ such that $\deg(f - dg) < \deg(f)$. If $f - dg \neq 0$ and $\deg(f - dg) \geq \deg(d)$ we choose a polynomial $h \in F[x]$ such that

$$\deg[f - d(g + h)] < \deg(f - dg).$$

Continuing this process as long as necessary, we ultimately obtain polynomials q, r satisfying (i) and (ii).

Suppose we also have $f = dq_1 + r_1$ where $\deg(r_1) < \deg(d)$. then $dq + r = dq_1 + r_1$ and $d(q - q_1) = r_1 - r$. If $q - a_1 \neq 0$, then $d(q - q_1) \neq 0$ and

$$\deg(d) + \deg(q - q_1) = \deg(r_1 - r).$$

But since the degree of $r_1 - r$ is less than the degree of d, this is impossible. Hence $q = a_1$ and then $r = r_1$. □

Definition Let d be a non-zero polynomial over the field F. If $f \in F[x]$, the preceding theorem shows that there is at most one polynomial $q \in F[x]$ such that $f = dq$. If such a q exists we say that d *divides* f, that f is *divisible* by d, and call q the quotient of f by d. We also write $q = f/d$.

Corollary 5.4.3. *Let $f \in F[x]$, and let $c \in F$. Then f is divisible by $x - c$ if and ony if $f(c) = 0$.*

Proof. By the theorem, $f = (x - c)q + r$ where r is a scalar polynomial. By Theorem 5.2.5,

$$f(c) = 0q(c) + r(c) = r(c).$$

Hence $r = 0$ if and only if $f(c) = 0$. □

Definition Let F be a field. An element $c \in F$ is said to be a *root* or a *zero* of a given polynomial $f \in F[x]$ provided $f(c) = 0$.

Corollary 5.4.4. *A polynomial $f \in F[x]$ of degree n has at most n roots in F.*

Proof. The result is obviously true for polynomials of degree 0 or 1. We assume it to be true for polynomials of degree $n - 1$. If a is a root of f, $f = (x - a)q$ where q has degree $n - 1$. Since $f(b) = 0$ if and only if $a = b$ or $q(b) = 0$, it follows by our induction hypothesis that f has at most n roots. □

Definition Let F be a field. An *ideal* in $F[x]$ is a subspace M of $F[x]$ such that fg belongs to M whenever $f \in F[x]$ and $g \in M$.

Example 5.4.5. *If F is a field and $d \in F[x]$, the set $M = dF[x]$ of all multiples df of d by arbitrary f in $F[x]$ is an ideal. This is because $d \in M$ (so M is nonempty), and it is easy to check that M is closed under addition and under multiplication by any element of $F[x]$. The ideal M is called the* principal ideal generated by *d and is denoted by $dF[x]$. If d is not the zero polynomial and its leading coefficient is a, then $d_1 = a^{-1}d$ is monic and $d_1F[x] = dF[x]$.*

Example 5.4.6. *Let $d_1, \ldots, d_n$ be a finite number of polynomials over F. Then the (vector space) sum M of the subspaces $d_iF[x]$ is a subspace and is also an ideal. M is the ideal generated by the polynomials $d_1, \ldots, d_n$.*

The following result is the main theorem on ideals in $F[x]$.

Theorem 5.4.7. *Let M be any non-zero ideal in $F[x]$. Then there is a unique monic polynomial $d \in F[x]$ such that M is the principal ideal $dF[x]$ generated by d.*

Proof. Among all nonzero polynomials in M there is (at least) one of minimal degree. Hence there must be a monic polynomial d of least degree in M. Suppose that f is any element of M. We can divide f by d and get a unique quotient and remainder: $f = qd + r$ where $\deg(r) < \deg(d)$. Then $r = f - qd \in M$, but $\deg(r)$ is less than the smallest degree of any nonzero polynomial in M. Hence $r = 0$. So $f = qd$. This shows that $M \subseteq dF[x]$. Clearly $d \in M$ implies $dF[x] \subseteq M$, so in fact $M = dF[x]$. If d_1 and d_2 are two monic polynomials in $F[x]$ for which $M = d_1F[x] = d_2F[x]$, then d_1 divides d_2 and d_2 divides d_1. Since they are monic, they clearly must be identical. □

Definition If $p_1, \ldots, p_k \in F[x]$ and not all of them are zero, then the monic generator d of the ideal $p_1F[x] + \cdots + p_kF[x]$ is called the *greatest comon divisor* (gcd) of $p_1, \cdots, p_k$. We say that the polynomials $p_1, \ldots, p_k$ are *relatively prime* if the greatest common divisor is 1, or equivalently if the ideal they generate is all of $F[x]$.

The exercises at the end of this chapter are to be considered an integral part of the chapter. You should study them all.

Definition The field F is called *algebraically closed* provided each polynomial in $F[x]$ that is irreducible over F has degree 1.

To say that F is algebraically closed means the every non-scalar irreducible monic polynomial over F is of the form $x - c$. So to say F is algebraically closed really means that each non-scalar polynomial f in $F[x]$ can be expressed in the form

$$f = c(x - c_1)^{n_1} \cdots (x - c_k)^{n_k}$$

where c is a scalar and $c_1, \ldots, c_k$ are distinct elements of F. It is also true that F is algebraically closed provided that each non-scalar polynomial over F has a root in F.

The field $\mathcal{R}$ of real numbers is not algebraically closed, since the polynomial $(x^2 + 1)$ is irreducible over $\mathcal{R}$ but not of degree 1. The *Fundamental Theorem of Algebra* states that the field $\mathcal{C}$ of complex numbers is algebraically closed. We shall prove this theorem later after we have introduced the concepts of determinant and of eigenvalues.

The Fundamental Theorem of Algebra also makes it clear what the possibilities are for the prime factorization of a polynomial with real coefficients. If f is a polynomial with real coefficients and c is a complex root of f, then the complex conjugate $\overline{c}$ is also a root of f. Therefore, those complex roots which are not real must occur in conjugate pairs, and the entire set of roots has the form $\{b_1, \ldots, b_r, c_1, \overline{c_1}, \ldots, c_k, \overline{c_k}\}$, where $b_1, \ldots, b_r$ are real and $c_1, \ldots, c_k$ are non-real complex numbers. Thus f factors as

$$f = c(x - b_1) \ldots (x - b_r) p_1 \cdots p_k$$

where p_i is the quadratic polynomial

$$p_i = (x - c_i)(x - \overline{c_i}).$$

These polynomials p_i have real coefficients. We see that every irreducible polynomial over the real number field has degree 1 or 2. Each polynomial over $\mathcal{R}$ is the product of certain linear factors given by the real roots of f and certain irreducible quadratic polynomials.

5.5 Exercises

1. (The Binomial Theorem) Let $\binom{m}{k} = \frac{m!}{k!(m-k)!}$ be the usual binomial coefficient. Let a, b be elements of any commutative ring. Then

$$(a + b)^m = \sum_{i=0}^{m} \binom{m}{k} a^{m-k} b^k.$$

Note that the binomial coefficient is an integer that may be reduced modulo any modulus p. It follows that even when the denominator of $\binom{m}{k}$ appears to be 0 in some ring of characteristic p, for example, this binomial coefficient can be interpreted modulo p. For example

$$\binom{6}{3} = 20 \equiv 2 \pmod 3,$$

even though 3! is zero modulo 3.

The *derivative* of the polynomial

$$f = c_0 + c_1 x + \cdots + c_n x^n$$

is the polynomial

$$f' = Df = c_1 + 2c_2x + \cdots + nc_nx^{n-1}.$$

Note: D is a linear operator on $F[x]$.

2. (Taylor's Formula). Let F be any field, let n be any positive integer, and let $f \in F$ have degree $m \leq n$. Then

$$f = \sum_{k=0}^{n} \frac{D^k(f)(c)}{k!}(x-c)^k.$$

Be sure to explain how to deal with the case where $k!$ is divisible by the characteristic of the field F.

If $f \in F[x]$ and $c \in F$, the *multiplicity* of c as a root of f is the largest positive integer r such that $(x-c)^r$ divides f.

3. Show that if the multiplicity of c as a root of f is $r \geq 2$, then the multiplicity of c as a root of f' is at least $r-1$.

4. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Let $M = \{f \in F[x] : f(A) = 0\}$. Show that M is a nonzero ideal. (Hint: consider the polynomial $f(x) = x^2 - (a+d)x + (ad-bc)$.)

Definition Let F be a field. A polynomial $f \in F[x]$ is said to be *reducible* over F provided there are polynomials $g, h \in F[x]$ with degree at least 1 for which $f = gh$. If f is not reducible over F, it is said to be *irreducible* over F. A polynomial $p(x) \in F[x]$ of degree at least 1 is said to be a *prime polynomial over* F provided whenever p divides a product gh of two polynomials in $F[x]$ then it has to divide at least one of g and h.

5. Show that a polynomial $p(x) \in F[x]$ with degree at least 1 is prime over F if and only if it is irreducible over F.

6. (The Primary Decomposition of f) If F is a field, a non-scalar monic polynomial in $F[x]$ can be factored as a product of monic primes in

$F[x]$ in one and, except for order, only one way. If $p_1, \ldots, p_k$ are the distinct monic primes occurring in this factorization of f, then

$$f = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k},$$

where n_i is the number of times the prime p_i occurs in this factorization. This decomposition is also clearly unique and is called the *primary decomposition* of f.

7. Let f be a non-scalar monic polynomial over the field F, and let

$$f = p_1^{n_1} \cdots p_k^{n_k}$$

be the prime factorization of f. For each j, $1 \leq j \leq k$, let

$$f_j = \frac{f}{p_j^{n_j}} = \prod_{i \neq j} p_i^{n_i}.$$

Then $f_1, \ldots, f_k$ are relatively prime.

8. Using the same notation as in the preceding problem, suppose that $f = p_1 \cdots p_k$ is a product of distinct non-scalar irreducible polynomials over F. So $f_j = f/p_j$. Show that

$$f' = p_1' f_1 + p_2' f_2 + \cdots + p_k' f_k.$$

9. Let $f \in F[x]$ have derivative f'. Then f is a product of distinct irreducible polynomials over F if and only if f and f' are relatively prime.

Chapter 6

Determinants

We assume that the reader has met the notion of a commutative ring K with 1. Our main goal in this chapter is to study the usual determinant function defined on the set of square matrices with entries from such a K. However, essentially nothing from the general theory of commutative rings with 1 will be used.

One of the main types of application of the notion of determinant is to determinants of matrices whose entries are polynomials in one or more indeterminates over a field F. So we might have $K = F[x]$, the ring of polynomials in the indeterminate x with coefficients from the field F. It is also quite useful to consider the theory of determinants over the ring $\mathcal{Z}$ of rational integers.

6.1 Determinant Functions

Throughout these notes K will be a commutative ring with identity. Then for each positive integer n we wish to assign to each $n \times n$ matrix over K a scalar (element of K) to be known as the *determinant* of the matrix. As soon as we have defined these terms we may say that the determinant function is n-linear alternating with value 1 at the identity matrix.

Definition Let D be a function which assigns to each $n \times n$ matrix A over K a scalar $D(A)$ in K. We say that D is *n-linear* provided that for each i, $1 \leq i \leq n$, D is a linear function of the ith row when the other $n-1$ rows are held fixed.

Perhaps this definition needs some clarification. If D is a function from

$M_{m,n}(K)$ into K, and if $\alpha_1, \ldots, \alpha_n$ are the rows of the matrix A, we also write

$$D(A) = D(\alpha_1, \ldots, \alpha_n),$$

that is, we think of D as a function of the rows of A. The statement that D is n-linear then means

$$\begin{aligned} D(\alpha_1, \ldots, c\alpha_i + \alpha_i', \ldots, \alpha_n) &= cD(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n) + \\ &+ D(\alpha_1, \ldots, \alpha_i', \ldots, \alpha_n). \end{aligned} \tag{6.1}$$

If we fix all rows except row i and regard D as a function of the ith row, it is often convenient to write $D(\alpha_i)$ for $D(A)$. Thus we may abbreviate Eq. 6.1 to

$$D(c\alpha_i + \alpha_i') = cD(\alpha_i) + D(\alpha_i'),$$

so long as it is clear what the meaning is.

In the following sometimes we use A_{ij} to denote the element in row i and column j of the matrix A, and sometimes we write $A(i, j)$.

Example 6.1.1. *Let $k_1, \ldots, k_n$ be positive integers, $1 \leq k_i \leq n$, and let a be any element of K. For each $n \times n$ matrix A over K, define*

$$D(A) = aA(1, k_1)A(2, k_2) \cdots A(n, k_n). \tag{6.2}$$

Then the function defined by Eq. 6.2 is n-linear. For, if we regard D as a function of the ith row of A, the others being fixed, we may write

$$D(\alpha_i) = A(i, k_i)b$$

where b is some fixed element of K. Let $\alpha_i' = (A_{i1}', \ldots, A_{in}')$. Then we have

$$\begin{aligned} D(c\alpha_i + \alpha_i') &= [cA(i, k_i) + A'(i, k_i)]b \\ &= cD(\alpha_i) + D(\alpha_i'). \end{aligned}$$

Thus D is a linear function of each of the rows of A.

A particular n-linear function of this type is just the product of the diagonal entries:

$$D(A) = A_{11}A_{22} \cdots A_{nn}.$$

Example 2. We find all 2-linear functions on 2×2 matrices over K. Let D be such a function. If we denote the rows of the 2×2 identity matrix by ϵ_1 and ϵ_2, then we have

$$D(A) = D(A_{11}\epsilon_1 + A_{12}\epsilon_2, A_{21}\epsilon_1 + A_{22}\epsilon_2).$$

Using the fact that D is 2-linear, we have

$$D(A) = A_{11}D(\epsilon_1, A_{21}\epsilon_1 + A_{22}\epsilon_2) + A_{12}D(\epsilon_2, A_{21}\epsilon_1 + A_{22}\epsilon_2) =$$

$$= A_{11}A_{21}D(\epsilon_1, \epsilon_1) + A_{11}A_{22}D(\epsilon_1, \epsilon_2) + A_{12}A_{21}D(\epsilon_2, \epsilon_1) + A_{12}A_{22}D(\epsilon_2, \epsilon_2).$$

This D is completely determined by the four scalars

$$D(\epsilon_1, \epsilon_1),\ D(\epsilon_1, \epsilon_2),\ D(\epsilon_2, \epsilon_1),\ D(\epsilon_2, \epsilon_2).$$

It is now routine to verify the following. If a, b, c, d are any four scalars in K and if we define

$$D(A) = A_{11}A_{21}a + A_{11}A_{22}b + A_{12}A_{21}c + A_{12}A_{22}d,$$

then D is a 2-linear function on 2×2 matrices over K and

$$D(\epsilon_1, \epsilon_1) = a, \quad D(\epsilon_1, \epsilon_2) = b$$
$$D(\epsilon_2, \epsilon_1) = c, \quad D(\epsilon_2, \epsilon_2) = d.$$

Lemma 6.1.2. *A linear combination of n-linear functions is n-linear.*

Proof. It suffices to prove that a linear combination of two n-linear functions is n-linear. Let D and E be n-linear functions. If a and b are elements of K the linear combination $aD + bE$ is defined by

$$(aD + bE)(A) = aD(A) + bE(A).$$

Hence, if we fix all rows except row i,

$$\begin{aligned}(aD + bE)(c\alpha_i + \alpha_i') &= aD(c\alpha_i + \alpha_i') + bE(c\alpha_i + \alpha_i') \\ &= acD(\alpha_i) + aD(\alpha_i') + bcE(\alpha_i) + bE(\alpha_i') \\ &= c(aD + bE)(\alpha_i) + (aD + bE)(\alpha_i').\end{aligned}$$

□

NOTE: If K is a field and V is the set of $n \times n$ matrices over K, the above lemma says the following. The set of n-linear functions on V is a subspace of the space of all functions from V into K.

Example 3. Let D be the function defined on 2×2 matrices over K by

$$D(A) = A_{11}A_{22} - A_{12}A_{21}. \tag{6.3}$$

This D is the sum of two functions of the type described in Example 1:

$$\begin{aligned} D &= D_1 + D_2 \\ D_1(A) &= A_{11}A_{22} \\ D_2(A) &= -A_{12}A_{21} \end{aligned} \tag{6.4}$$

Most readers will recognize this D as the "usual" determinant function and will recall that it satisfies several additional properties, such as the following one.

6.1.3 n-Linear Alternating Functions

Definition: Let D be an n-linear function. We say D is *alternating* provided $D(A) = 0$ whenever two rows of A are equal.

Lemma 6.1.4. *Let D be an n-linear alternating function, and let A be $n \times n$. If A' is obtained from A by interchanging two rows of A, then $D(A') = -D(A)$.*

Proof. If the ith row of A is α and the jth row of A is β, $i \neq j$, and all other rows are being held constant, we write $D(\alpha, \beta)$ in place of $D(A)$.

$$D(\alpha + \beta, \alpha + \beta) = D(\alpha, \alpha) + D(\alpha, \beta) + D(\beta, \alpha) + D(\beta, \beta).$$

By hypothesis, $D(\alpha + \beta, \alpha + \beta) = D(\alpha, \alpha) = D(\beta, \beta) = 0$. So

$$0 = D(\alpha, \beta) + D(\beta, \alpha).$$

□

If we assume that D is n-linear and has the property that $D(A') = -D(A)$ when A' is obtained from A by interchanging any two rows of A, then if A has two equal rows, clearly $D(A) = -D(A)$. If the characteristic of the ring K is odd or zero, then this forces $D(A) = 0$, so D is alternating. But, for example, if K is an integral domain with characteristic 2, this is clearly not the case.

6.1.5 A Determinant Function - The Laplace Expansion

Definition: Let K be a commutative ring with 1, and let n be a positive integer. Suppose D is a function from $n \times n$ matrices over K into K. We say that D is a *determinant function* if D is n-linear, alternating and $D(I) = 1$.

It is clear that there is a unique determinant function on 1×1 matrices, and we are now in a position to handle the 2×2 case. It should be clear that the function given in Example 3. is a determinant function. Furthermore, the formulas exhibited in Example 2. make it easy to see that the D given in Example 3. is the unique determinant function.

Lemma 6.1.6. *Let D be an n-linear function on $n \times n$ matrices over K. Suppose D has the property that $D(A) = 0$ when any two adjacent rows of A are equal. Then D is alternating.*

Proof. Let B be obtained by interchanging rows i and j of A, where $i < j$. We can obtain B from A by a succession of interchanges of pairs of adjacent rows. We begin by interchanging row i with row $i+1$ and continue until the rows are in the order

$$\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_j, \alpha_i, \alpha_{j+1}, \ldots, \alpha_n.$$

This requires $k = j - i$ interchanges of adjacent rows. We now move α_j to the ith position using $(k-1)$ interchanges of adjacent rows. We have thus obtained B from A by $2k-1$ interchanges of adjacent rows. Thus by Lemma 6.1.4,

$$D(B) = -D(A).$$

Suppose A is any $n \times n$ matrix with two equal rows, say $\alpha_i = \alpha_j$ with $i < j$. If $j = i + 1$, then A has two equal and adjacent rows, so $D(A) = 0$.

If $j > i+1$, we intechange α_{i+1} and α_j and the resulting matrix B has two equal and adjacent rows, so $D(B) = 0$. On the other hand, $D(B) = -D(A)$, hence $D(A) = 0$. □

Lemma 6.1.7. *Let K be a commutative ring with 1 and let D be an alternating n-linear function on $n \times n$ matrices over K. Then*

(a) $D(A) = 0$ if one of the rows of A is 0.

(b) $D(B) = D(A)$ if B is obtained from A by adding a scalar multiple of one row of A to a different row of A.

Proof. For part (a), suppose the ith row α_i is a zero row. Using the linearity of D in the ith row of A says $D(\alpha_i + \alpha_i) = D(\alpha_i) + D(\alpha_i)$, which forces $D(A) = D(\alpha_i) = 0$. For part (b), if $i \neq j$, write $D(A) = D(\alpha_i, \alpha_j)$, with all rows other than the ith one held fixed. $D(B) = D(\alpha_i + c\alpha_j, \alpha_j) = D(\alpha_i, \alpha_j) + cD(\alpha_j, \alpha_j) = D(A) + 0$. □

Definition: If $n > 1$ and A is an $n \times n$ matrix over K, we let $A(i|j)$ denote the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of A. If D is an $(n-1)$-linear function and A is an $n \times n$ matrix, we put $D_{ij}(A) = D[A(i|j)]$.

Theorem 6.1.8. *Let $n > 1$ and let D be an alternating $(n-1)$-linear function on $(n-1) \times (n-1)$ matrices over K. For each j, $1 \leq j \leq n$, the function E_j defined by*

$$E_j(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} D_{ij}(A) \tag{6.5}$$

is an alternating n-linear function on $n \times n$ matrices A. If D is a determinant function, so is each E_j.

Proof. If A is an $n \times n$ matrix, $D_{ij}(A)$ is independent of the ith row of A. Since D is $(n-1)$-linear, it is clear that D_{ij} is linear as a function of any row except row i. Therefore $A_{ij}D_{ij}(A)$ is an n-linear function of A. Hence E_j is n-linear by Lemma 6.1.2. To prove the E_j is alternating it will suffice to show that $E_j(A) = 0$ whenever A has two equal and adjacent rows. Suppose $\alpha_k = \alpha_{k+1}$. If $i \neq k$ and $i \neq k+1$, the matrix $A(i|j)$ has two equal rows, and thus $D_{ij}(A) = 0$. Therefore,

$$E_j(A) = (-1)^{k+j} A_{kj} D_{kj}(A) + (-1)^{k+1+j} A_{(k+1)j} D_{(k+1)j}(A).$$

Since $\alpha_k = \alpha_{k+1}$,

$$A_{kj} = A_{(k+1)j} \text{ and } A(k|j) = A(k+1|j).$$

Clearly then $E_j(A) = 0$.

Now suppose D is a determinant function. If $I^{(n)}$ is the $n \times n$ identity matrix, then $I^{(n)}(j|j)$ is the $(n-1) \times (n-1)$ identity matrix $I^{(n-1)}$. Since $I^{(n)}_{ij} = \delta_{ij}$, it follows from Eq. 6.5 that

$$E_j(I^{(n)}) = D(I^{(n-1)}). \tag{6.6}$$

Now $D(I^{(n-1)}) = 1$, so that $E_j(I^{(n)}) = 1$ and E_j is a determinant function. □

We emphasize that this last Theorem (together with a simple induction argument) shows that if K is a commutative ring with identity and $n \geq 1$, then there exists at least one determinant function on $K^{n\times n}$. In the next section we will show that there is only one determinant function. The determinant function E_j is referred to as the Laplace expansion of the determinant along the jth column. There is a similar Laplace expansion of the determinant along the ith row of A which will eventually show up as an easy corollary.

6.2 Permutations & Uniqueness of Determinants

6.2.1 A Formula for the Determinant

Suppose that D is an alternating n-linear function on $n \times n$ matrices over K. Let A be an $n \times n$ matrix over K with rows $\alpha_1, \alpha_2, \ldots, \alpha_n$. If we denote the rows of the $n \times n$ identity matrix over K by $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, then

$$\alpha_i = \sum_{j=1}^{n} A_{ij}\epsilon_j, \quad 1 \leq i \leq n. \tag{6.7}$$

Hence

$$\begin{aligned} D(A) &= D\left(\sum_j A_{1j}\epsilon_j, \alpha_2, \ldots, \alpha_n\right) \\ &= \sum_j A_{1j} D(\epsilon_j, \alpha_2, \ldots, \alpha_n). \end{aligned}$$

If we now replace α_2 with $\sum_k A_{2k}\epsilon_k$, we see that

$$D(A) = \sum_{j,k} A_{1j}A_{2k} D(\epsilon_j, \epsilon_k, \ldots, \alpha_n).$$

In this expression replace α_3 by $\sum_l A_{3l}\epsilon_l$, etc. We finally obtain

$$D(A) = \sum_{k_1,k_2,\ldots,k_n} A_{1k_1}A_{2k_2}\cdots A_{nk_n} D(\epsilon_{k_1}, \ldots, \epsilon_{k_n}). \tag{6.8}$$

Here the sum is over all sequences $(k_1, k_2, \ldots, k_n)$ of positive integers not exceeding n. This shows that D is a finite sum of functions of the type described by Eq. 6.2. Note that Eq. 6.8 is a consequence just of the assumption that D is n-linear, and that a special case was obtained earlier in Example 2. Since D is alternating,

$$D(\epsilon_{k_1}, \epsilon_{k_2}, \ldots, \epsilon_{k_n}) = 0$$

whenever two of the indices k_i are equal. A sequence $(k_1, k_2, \ldots, k_n)$ of positive integers not exceeding n, with the property that no two of the k_i are equal, is called a *permutation of degree* n. In Eq. 6.8 we need therefore sum only over those sequences which are permutations of degree n.

A permutation of degree n may be defined as a one-to-one function from the set $\{1, 2, \ldots, n\}$ onto itself. Such a function σ corresponds to the n-tuple $(\sigma 1, \sigma 2, \ldots, \sigma n)$ and is thus simply a rule for ordering $1, 2, \ldots, n$ in some well-defined way.

If D is an alternating n-linear function and A is a $n \times n$ matrix over K, we then have

$$D(A) = \sum_\sigma A_{1(\sigma 1)}\cdots A_{n(\sigma n)} D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) \tag{6.9}$$

where the sum is extended over the distinct permutations σ of degree n.

Next we shall show that

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = \pm D(\epsilon_1, \ldots, \epsilon_n) \tag{6.10}$$

where the sign $\pm$ depends only on the permutation σ. The reason for this is as follows. The sequence $(\sigma 1, \sigma 2, \ldots, \sigma n)$ can be obtained from the sequence $(1, 2, \ldots, n)$ by a finite number of interchanges of pairs of elements. For example, if $\sigma 1 \neq 1$, we can transpose 1 and $\sigma 1$, obtaining $(\sigma 1, \ldots, 1, \ldots)$. Proceeding in this way we shall arrive at the sequence $(\sigma 1, \ldots, \sigma n))$ after n or fewer such interchanges of pairs. Since D is alternating, the sign of its value changes each time that we interchange two of the rows ϵ_i and ϵ_j. Thus, if we pass from $(1, 2, \ldots, n)$ to $(\sigma 1, \sigma 2, \ldots, \sigma n)$ by means of m interchanges of pairs (i, j), we shall have

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = (-1)^m D(\epsilon_1, \ldots, \epsilon_n).$$

In particular, if D is a determinant function

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = (-1)^m, \tag{6.11}$$

where m depends only on σ, not on D. Thus all determinant functions assign the same value to the matrix with rows $\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}$, and this value is either 1 or -1.

A basic fact about permutations is the following: if σ is a permutation of degree n, one can pass from the sequence $(1, 2, \ldots, n)$ to the sequence $(\sigma 1, \sigma 2, \ldots, \sigma n)$ by a succession of interchanges of pairs, and this can be done in a variety of ways. However, no matter how it is done, the number of interchanges used is either always even or always odd. The permutation is then called *even* or *odd*, respectively. One defines the *sign* of a permutation by

$$\text{sgn } \sigma = \begin{cases} 1, & \text{if } \sigma \text{ is even;} \\ -1, & \text{if } \sigma \text{ is odd.} \end{cases}$$

We shall establish this basic properties of permutations below from what we already know about determinant functions. However, for the moment let us assume this property. Then the integer m occurring in Eq. 6.11 is always even if σ is an even permutation, and is always odd if σ is an odd permutation. For any alternating n-linear function D we then have

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = (\text{sgn } \sigma) D(\epsilon_1, \ldots, \epsilon_n),$$

and using Eq. 6.9 we obtain

$$D(A) = \left[\sum_{\sigma}(\text{sgn } \sigma)A_{1(\sigma 1)} \cdots A_{n(\sigma n)}\right] D(I). \tag{6.12}$$

From Eq. 6.12 we see that there is precisely one determinant function on $n \times n$ matrices over K. If we denote this function by det, it is given by

$$\det(A) = \sum_{\sigma}(\text{sgn } \sigma)A_{1(\sigma 1)}A_{2(\sigma 2)} \cdots A_{n(\sigma n)}, \tag{6.13}$$

the sum being extended over the distinct permutations σ of degree n. We can formally summarize this as follows.

Theorem 6.2.2. *Let K be a commutative ring with 1 and let n be a positive integer. There is precisely one determinant function on the set of $n \times n$ matrices over K and it is the function* det *defined by Eq. 6.13. If D is any alternating n-linear function on $M_n(K)$, then for each $n \times n$ matrix A,*

$$D(A) = (det\ A)D(I).$$

This is the theorem we have been working towards, but we have left a gap in the proof. That gap is the proof that for a given permutation σ, when we pass from $(1, 2, \ldots, n)$ to $(\sigma 1, \ldots, \sigma n)$ by interchanging pairs, the number of interchanges is always even or always odd. This basic combintaorial fact can be proved without any reference to determinants. However, we now point out how it follows from the *existence* of a determinant function on $n \times n$ matrices.

Let K be the ring of rational integers. Let D be a determinant function on the $n \times n$ matrices over K. Let σ be a permutation of degree n, and suppose we pass from $(1, 2, \ldots, n)$ to $(\sigma 1, \ldots, \sigma n)$ by m interchanges of pairs (i, j), $i \neq j$. As we showed in Eq. 6.11

$$(-1)^m = D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}),$$

that is, the nuber $(-1)^m$ must be the value of D on the matrix with rows $\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}$. If

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = 1,$$

then m must be even. If

$$D(\epsilon_{\sigma 1}, \ldots, \epsilon_{\sigma n}) = -1,$$

then m must be odd.

From the point of view of products of permutations, the basic property of the sign of a permutation is that

$$\text{sgn } (\sigma\tau) = (\text{sgn } \sigma)(\text{sgn } \tau). \tag{6.14}$$

This result also follows from the theory of determinants (but is well known in the theory of the symmetric group independet of any determinant theory). In fact, it is an easy corollary of the following theorem.

Theorem 6.2.3. *Let K be a commutative ring with identity, and let A and B be $n \times n$ matrices over K. Then*

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

Proof. Let B be a fixed $n \times n$ matrix over K, and for each $n \times n$ matrix A define $D(A) = \text{ det } (AB)$. If we denote the rows of A by $\alpha_1, \ldots, \alpha_n$, then

$$D(\alpha_1, \ldots, \alpha_n) = \text{ det } (\alpha_1 B, \ldots, \alpha_n B).$$

Here $\alpha_j B$ denotes the $1 \times n$ matrix which is the product of the $1 \times n$ matrix α_j and the $n \times n$ matrix B. Since

$$(c\alpha_i + \alpha_i')B = c\alpha_i B + \alpha_i' B$$

and det is n-linear, it is easy to see that D is n-linear. If $\alpha_i = \alpha_j$, then $\alpha_i B = \alpha_j B$, and since det is alternating,

$$D(\alpha_1, \ldots, \alpha_n) = 0.$$

Hence, D is alternating. So D is an alternating n-linear function, and by Theorem 6.2.2

$$D(A) = (\text{det } A)D(I).$$

But $D(I) = \text{ det } (IB) = \text{ det } B$, so

$$\text{det } (AB) = D(A) = (\text{det } A)(\text{det } B).$$

□

6.3 Additional Properties of Determinants

Several well-known properties of the determinant function are now easy consequences of results we have already obtained, Eq. 6.13 and Theorem 6.2.3, for example. We give a few of these, proving some and leaving the others as rather routine exercises.

6.3.1 If A is a unit in $M_n(K)$, then det(A) is a unit in K.

If A is an invertible $n \times n$ matrix over a commutative ring K with 1, then det (A) is a unit in the ring K.

6.3.2 Triangular Matrices

If the square matrix A over K is upper or lower triangular, then det (A) is the product of the diagonal entries of A.

6.3.3 Transposes

If A^T is the transpose of the square matrix A, then
det $(A^T) =$ det (A).

Proof. If σ is a permutation of degree n,

$$(A^T)_{i(\sigma i)} = A_{(\sigma i)i}.$$

Hence

$$\text{det } (A^T) = \sum_{\sigma} (\text{sgn } \sigma) A_{(\sigma 1)1} \cdots A_{(\sigma n)n}.$$

When $i = \sigma^{-1} j$, $A_{(\sigma i)i} = A_{j(\sigma^{-1} j)}$. Thus

$$A_{(\sigma 1)1} \cdots A_{(\sigma n)n} = A_{1(\sigma 1)} \cdots A_{n(\sigma n)}.$$

Since $\sigma\sigma^{-1}$ is the identity permutation,

$$(\text{sgn } \sigma)(\text{sgn } \sigma^{-1}) = 1, \text{ so sgn } (\sigma^{-1}) = \text{sgn } (\sigma).$$

Furthermore, as σ varies over all permutations of degree n, so does σ^{-1}. Therefore,

$$\det\ (A^T) = \sum_{\sigma} (\text{sgn}\ \sigma^{-1}) A_{1(\sigma^{-1}1)} \cdots A_{n(\sigma^{-1}n)} = \det\ (A).$$

□

6.3.4 Elementary Row Operations

If B is obtained from A by adding a multiple of one row of A to another (or a multiple of one column to another), then $\det(A) = \det(B)$. If $B = cA$, then $\det\ (B) = c^n \det\ (A)$.

6.3.5 Triangular Block Form

Suppose an $n \times n$ matrix A is given in block form

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

where B is an $r \times r$ matrix, E is an $s \times s$ matrix, C is $r \times s$, and 0 denotes the $s \times r$ zero matrix. Then

$$\det \begin{pmatrix} B & C \\ 0 & E \end{pmatrix} = (\det\ B)(\det\ E). \tag{6.15}$$

Proof. To prove this, define

$$D(B, C, E) = \det \begin{pmatrix} B & C \\ 0 & E \end{pmatrix}.$$

If we fix B and C, then D is alternating and s-linear as a function of the rows of E. Hence by Theorem 6.2.2

$$D(B, C, E) = (\det\ E) D(B, C, I),$$

where I is the $s \times s$ identity matrix. By subtracting multiples of the rows of I from the rows of B and using the result of 6.3.4, we obtain

$$D(B, C, I) = D(B, 0, I).$$

Now $D(B, 0, I)$ is clearly alternating and r-linear as a function of the rows of B. Thus

$$D(B, 0, I) = (\det\ B)D(I, 0, I) = 1.$$

Hence

$$\begin{aligned} D(B, C, E) &= (\det\ E)D(B, C, I) \\ &= (\det\ E)D(B, 0, I) \\ &= (\det\ E)(\det\ B). \end{aligned}$$

□

By taking transposes we obtain

$$\det \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} = (\det\ A)(\det\ C).$$

It is now easy to see that this result generalizes immediately to the case where A is in upper (or lower) block triangular form.

6.3.6 The Classical Adjoint and the Laplace Expansion

Since the determinant function is unique and $\det(A) = \det(A^T)$, we know that for each fixed column index j,

$$\det\ (A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det\ A(i|j), \tag{6.16}$$

and for each row index i,

$$\det\ (A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det\ A(i|j). \tag{6.17}$$

As we mentioned earlier, the formulas of Eqs. 6.16 and 6.17 are known as the Laplace expansion of the determinant in terms of columns, respectively, rows. Later we will present a more general version of the Laplace expansion.

The scalar $(-1)^{i+j} \det\ A(i|j)$ is usually called the i, j *cofactor* of A or the cofacor of the i, j entry of A. The above formulas for $\det\ (A)$ are called the

expansion of det (A) by cofactors of the jth column (sometimes the expansion by minors of the jth column), or respectively, the expansion of det (A) by cofactors of the ith row (sometimes the expansion by minors of the ith row). If we set

$$C_{ij} = (-1)^{i+j}\text{det } A(i|j),$$

then the formula in Eq. 6.16 says that for each j,

$$\text{det } (A) = \sum_{i=1}^{n} A_{ij}C_{ij},$$

where the cofactor C_{ij} is $(-1)^{i+j}$ times the determinant of the $(n-1)\times(n-1)$ matrix obtained by deleting the ith row and jth column of A.

Similarly, for each fixed row index i,

$$\text{det } (A) = \sum_{j=1}^{n} A_{ij}C_{ij}.$$

If $j \neq k$, then

$$\sum_{i=1}^{n} A_{ik}C_{ij} = 0.$$

To see this, replace the jth column of A by its kth column, and call the resulting matrix B. Then B has two equal columns and so $\det(B) = 0$. Since $B(i|j) = A(i|j)$, we have

$$\begin{aligned} 0 &= \text{det } (B) \\ &= \sum_{i=1}^{n}(-1)^{i+j}B_{ij}\text{det } (B(i|j)) \\ &= \sum_{i=1}^{n}(-1)^{i+j}A_{ik}\text{det } (A(i|j)) \\ &= \sum_{i=1}^{n} A_{ik}C_{ij}. \end{aligned}$$

These properties of the cofactors can be summarized by

$$\sum_{i=1}^{n} A_{ik}C_{ij} = \delta_{jk}\text{det } (A). \tag{6.18}$$

The $n \times n$ matrix adj A , which is the transpose of the matrix of cofactors of A is called the *classical adjoint* of A. Thus

$$(\text{adj } A)_{ij} = C_{ji} = (-1)^{i+j}\text{det } (A(j|i)). \tag{6.19}$$

These last two formulas can be summarized in the matrix equation

$$(\text{adj } A)A = (\text{det } (A))I. \tag{6.20}$$

We wish to see that $A(\text{adj } A) = (\text{det } A)I$ also. Since $A^T(i|j) = (A(j|i))^T$, we have

$$(-1)^{i+j}\text{det } (A^T(i|j)) = (-1)^{i+j}\text{det } (A(j|i)),$$

which simply says that the i, j cofactor of A^T is the j, i cofactor of A. Thus

$$\text{adj } (A^T) = (\text{adj } A)^T. \tag{6.21}$$

Applying this last equation to A^T, we have

$$(\text{adj } A^T)A^T = (\text{det } (A^T))I = (\text{det } A)I.$$

Transposing, we obtain

$$A(\text{adj } A^T)^T = (\text{det } (A))I.$$

Using Eq. 6.21 we have what we want:

$$A(\text{adj } A) = (\text{det } (A))I. \tag{6.22}$$

An almost immediate corollary of the previous paragraphs is the following:

Theorem 6.3.7. *Let A be an $n \times n$ matrix over K. Then A is invertible over K if and only if det (A) is invertible in K. When A is invertible, the unique inverse for A is*

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

In particular, an $n \times n$ matrix over a field is invertible if and only if its determinant is different from zero.

NOTE: This determinant criterion for invertibility proves that an $n \times n$ matrix with either a left or right inverse is invertible.

NOTE: The reader should think about the consequences of Theorem 6.3.7 in case K is the ring $F[x]$ of polynomials over a field F, or in case K is the ring of rational integers.

6.3.8 Characteristic Polynomial of a Linear Map

If P is also an $n \times n$ invertible matrix, then because K is commutative and det is multiplicative, it is immediate that

$$\det\ (P^{-1}AP) = \det\ (A). \tag{6.23}$$

This means that if K is actually a field, if V is an n-dimensional vector space over K, if $T : V \to V$ is any linear map, and if $\mathcal{B}$ is any basis of V, then we may unambiguously define the *characteristic polynomial* $c_T(x)$ of T to be

$$c_T(x) = det(xI - [T]_{\mathcal{B}}).$$

This is because if A and B are two matrices that represent the same linear transformation with respect to some bases of V, then by Eq. 6.23 and Theorem 4.7.1

$$det(xI - A) = det(xI - B).$$

6.3.9 Coefficients of the Characteristic Polynomial

Theorem 6.3.10. *Let K be a commutative ring with 1, and let A be an $n \times n$ matrix over K. The characteristic polynomial of A is given by*

$$f(x) = det(xI - A) = \sum_{i=0}^{n} c_i x^{n-i} \tag{6.24}$$

where $c_0 = 1$, and for $1 \leq i \leq n$, $c_i = \sum det(B)$, where B ranges over all the $i \times i$ principal submatrices of $-A$.

For an $n \times n$ matrix A, the *trace of* A is defined to be

$$\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

Note: Putting $i = 1$ yields the fact that the coefficient of x^{n-1} is $-\sum_{i=1}^{n} A_{ii} = -\mathrm{tr}(A)$, and putting $i = n$ says that the constant term is $(-1)^n \mathrm{det}(A)$.

Proof. Clearly $det(xI - A)$ is a polynomial of degree n which is monic, i.e., $c_0 = 1$, and and with constant term $det(-A) = (-1)^n det(A)$. Suppose $1 \leq i \leq n-1$ and consider the coefficient c_i of x^{n-i} in the polynomial $det(xI - A)$. Recall that in general, if $D = (d_{ij})$ is an $n \times n$ matrix over a commutative ring with 1, then

$$det(D) = \sum_{\pi \in \mathcal{S}_n} (-1)^{sgn(\pi)} \cdot d_{1,\pi(1)} d_{2,\pi(2)} \cdots d_{n,\pi(n)}.$$

So to get a term of degree $n - i$ in $det(xI - A) = \sum_{\pi \in \mathcal{S}_n} (-1)^{sgn(\pi)} (xI - A)_{1,\pi(1)} \cdots (xI - A)_{n,\pi(n)}$ we first select $n-i$ indices $j_1, \ldots, j_{n-i}$, with complementary indices $k_1, \ldots, k_i$. Then in expanding the product $(xI - A)_{1,\pi(1)} \cdots (xI - A)_{n,\pi(n)}$ when π fixes $j_1, \ldots, j_{n-i}$, we select the term x from the factors $(xI - A)_{j_1,j_1}, \ldots, (xI - A)_{j_{n-i}j_{n-i}}$, and the terms $(-A)_{k_1,\pi(k_1)}, \ldots, (-A)_{k_i,\pi(k_i)}$ otherwise. So if $A(k_1, \ldots, k_i)$ is the principal submatrix of A indexed by rows and columns $k_1, \ldots, k_i$, then $det(-A(k_1, \ldots, k_i))$ is the associated contribution to the coefficient of x^{n-i}. It follows that $c_i = \sum det(B)$ where B ranges over all the principal $i \times i$ submatrices of $-A$. □

Suppose the permutation $\pi \in \mathcal{S}_n$ consists of k permutation cycles of sizes $l_1, \ldots, l_k$, respectively, where $\sum l_i = n$. Then $sgn(\pi)$ can be computed by

$$sgn(\pi) = (-1)^{l_1 - 1 + l_2 - 1 + \cdots l_k - 1} = (-1)^{n-k} = (-1)^n (-1)^k.$$

We record this formally as:

$$sgn(\pi) = (-1)^n (-1)^k \text{ if } \pi \in \mathcal{S}_n \text{ is the product of } k \text{ disjoint cycles.} \quad (6.25)$$

6.3.11 The Companion Matrix of a Polynomial

In this section K is a field and $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$. Define the *companion matrix* $C(f(x))$ by

$$C(f(x)) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

The main facts about $C(f(x))$ are in the next result.

Theorem 6.3.12.

$$det\,(xI_n - C(f(x))) = f(x)$$

is both the minimal and characteristic polynomial of $C(f(x))$.

Proof. First we establish that $f(x) = det(xI_n - C(f(x)))$. This result is clear if $n = 1$ and we proceed by induction. Suppose that $n > 1$ and compute the determinant by cofactor expansion along the first row, applying the induction hypothesis to the first summand.

$$det(xI_n - C(f(x))) = det \begin{pmatrix} x & 0 & \cdots & 0 & 0 & a_0 \\ -1 & x & \cdots & 0 & 0 & a_1 \\ 0 & -1 & \cdots & 0 & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & x & a_n - 2 \\ 0 & 0 & \cdots & 0 & -1 & x + a_{n-1} \end{pmatrix}.$$

$$= x\ det \begin{pmatrix} x & \cdots & 0 & 0 & a_1 \\ -1 & \cdots & 0 & 0 & a_2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & -1 & x & a_{n-2} \\ 0 & \cdots & 0 & -1 & x + a_{n-1} \end{pmatrix}$$

$$+a_0(-1)^{n+1}det\begin{pmatrix} -1 & x & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & x \\ 0 & 0 & \cdots & - & -1 \end{pmatrix}$$

$$= x(x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_1) + a_0(-1)^{n+1}(-1)^{n-1}$$

$$= x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = f(x).$$

This shows that $f(x)$ is the characteristic polynomial of $C(f(x))$.

Now let T be the linear operator on K^n whose matrix with respect to the standard basis $\mathcal{S} = (e_1, e_2, \ldots, e_n)$ is $C(f(x))$. Then $Te_1 = e_2$, $T^2e_1 = Te_2 = e_3$, $\ldots$, $T^je_1 = T(e_j) = e_{j+1}$ for $1 \leq j \leq n-1$, and $Te_n = -a_0e_1 - a_1e_2 - \cdots - a_{n-1}e_n$, so

$$(T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a + 0I)e_1 = \overline{0}.$$

Also

$$(T^n + \cdots + a_1T + a_0I)e_{j+1} = (T^n + \cdots + a_1T + a_0I)T^je_1$$

$$= T^j(T^n + \cdots + a_1T + a_0I)e_1 = \overline{0}.$$

It follows that $f(T)$ must be the zero operator. On the other hand, $(e_1, Te_1, \ldots, T^{n-1}e_1)$ is a linearly independent list, so that no nonzero polynomial in T with degree less than n can be the zero operator. Then since $f(x)$ is monic it must be that $f(x)$ is also the minimal polynomial for T and hence for $C(f(x))$. □

6.3.13 The Cayley-Hamilton Theorem

Let $\dim(V) = n$ and let $T \in \mathcal{L}(V)$. If f is the characteristic polynomial for T, then $f(T) = 0$. This is equivalent to saying that the minimal polynomial for T divides the characteristic polynomial for T.

Proof. This proof is an illuminating and fairly sophisticated application of the general theory of determinants developed above.

Let K be the commutative ring with identity consisting of all polynomials in T. Actually, K is a commutative algebra with identity over the scalar

field F. Choose a basis $\mathcal{B} = (v_1, \ldots, v_n)$ for V and let A be the matrix which represents T in the given basis. Then

$$T(v_j) = \sum_{i=1}^{n} A_{ij} v_i, \quad 1 \leq j \leq n.$$

These equations may be written in the equivalent form

$$\sum_{i=1}^{n} (\delta_{ij} T - A_{ij} I) v_i, \quad 1 \leq j \leq n.$$

Let $B \in M_n(K)$ be the matrix with entries

$$B_{ij} = \delta_{ij} T - A_{ji} I.$$

Note the interchanging of the i and j in the subscript on A. Let $f(x) = \det(xI - A) = \det\left(xI - A^T\right)$. Then $f(T) = \det(B)$. Our goal is to show that $f(T) = 0$. In order that $f(T)$ be the zero operator, it is necessary and sufficient that $\det(B)(v_k) = \vec{0}$ for $1 \leq k \leq n$. By the definition of B, the vectors $v_1, \ldots, v_n$ satisfy the equations

$$\vec{0} = \sum_{i=1}^{n} (\delta_{ij} T - A_{ij} I)(v_i) = \sum_{i=1}^{n} B_{ji} v_i, \ 1 \leq j \leq n. \tag{6.26}$$

Let $\tilde{B}$ be the classical adjoint of B, so that $\tilde{B}B = B\tilde{B} = \det(B) I$. Note that $\tilde{B}$ also has entries that are polynomials in the operator T. Let $\tilde{B}_{kj}$ operate on the right side of Eq. 6.26 to obtain

$$\vec{0} = \tilde{B}_{kj} \sum_{i=1}^{n} B_{ji}(v_i) = \sum_{i=1}^{n} (\tilde{b}_{kj} B_{ji})(v_i).$$

So summing over j we have

$$\vec{0} = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \tilde{B}_{kj} B_{ji} \right) (v_i) = \sum_{i=1}^{n} \left(\delta_{ki} \det(B) \right) (v_i) = \det(B)(v_k).$$

□

At this point we know that each irreducible factor of the minimal polynomial of T is also a factor of the characteristic polynomial of T. A converse is also true: Each irreducible factor of the characteristic polynomial of T is also a factor of the minimal polynomial of T. However, we are not yet ready to give a proof of this fact.

6.3.14 Cramer's Rule

We now discuss **Cramer's rule** for solving systems of linear equations. Suppose A is an $n \times n$ matrix over the field F and we wish to solve the system of linear equations $AX = Y$ for some given n-tuple $(y_1, \ldots, y_n)$. If $AX = Y$, then

$$(\text{adj } A)AX = (\text{adj } A)Y$$

implying

$$(\det A)X = (\text{adj } A)Y.$$

Thus

$$\begin{aligned}(\det A)x_j &= \sum_{i=1}^{n}(\text{adj } A)_{ji}y_i \\ &= \sum_{i=1}^{n}(-1)^{i+j}y_i \det A(i|j).\end{aligned}$$

This last expression is the determinant of the $n \times n$ matrix obtained by replacing the jth column of A by Y. If $\det A = 0$, all this tells us nothing. But if $\det A \neq 0$, we have *Cramer's rule*:

Let A be an $n \times n$ matrix over the field F such that $\det A \neq 0$. If $y_1, \ldots, y_n$ are any scalars in F, the unique solution $X = A^{-1}Y$ of the system of equations $AX = Y$ is given by

$$x_j = \frac{\det B_j}{\det A}, \quad j = 1, \ldots, n,$$

where B_j is the $n \times n$ matrix obtained from A by replacing the jth column of A by Y.

6.4 Deeper Results with Some Applications*

The remainder of this chapter may be omitted without loss of continuity.

In general we continue to let K be a commutative ring with 1. Here $Mat_n(K)$ denotes the ring of $n \times n$ matrices over K, and $|A|$ denotes the element of K that is the determinant of A.

6.4.1 Block Matrices whose Blocks Commute*

We can regard a $k \times k$ matrix $M = (A^{(i,j)})$ over $Mat_n(K)$ as a *block matrix*, a matrix that has been partitioned into k^2 submatrices (*blocks*) over K, each of size $n \times n$. When M is regarded in this way, we denote its determinant in K by $|M|$. We use the symbol $D(M)$ for the determinant of M viewed as a $k \times k$ matrix over $Mat_n(K)$. It is important to realize that $D(M)$ is an $n \times n$ matrix.

Theorem. Assume that M is a $k \times k$ block matrix of $n \times n$ blocks $A^{(i,j)}$ over K that pairwise commute. Then

$$|M| = |D(M)| = \left| \sum_{\sigma \in \mathcal{S}_k} (\text{sgn } \sigma) A^{(1,\sigma(1))} A^{(2,\sigma(2))} \cdots A^{(k,\sigma(k))} \right| . \tag{6.27}$$

Here $\mathcal{S}_k$ is the symmetric group on k symbols, so the summation is the usual one that appears in a formula for the determinant. The first proof of this result to come to our attention is the one in N. Jacobson, *Lectures in Abstract algebra*, Vol. III – *Theory of Fields and Galois Theory*, D. Van Nostrand Co., Inc., 1964, pp 67 – 70. The proof we give now is from I. Kovacs, D. S. Silver, and Susan G. Williams, Determinants of Commuting-Block Matrices, *Amer. Math. Monthly*, Vol. 106, Number 10, December 1999, pp. 950 – 952.

Proof. We use induction on k. The case $k = 1$ is evident. We suppose that Eq. 6.27 is true for $k-1$ and then prove it for k. Observe that the following matrix equation holds:

$$\begin{pmatrix} I & 0 & \cdots & 0 \\ -A^{(2,1)} & I & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ -A^{(k,1)} & 0 & \cdots & I \end{pmatrix} \begin{pmatrix} I & & \cdots & 0 \\ 0 & A^{(1,1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A^{(1,1)} \end{pmatrix} M = \begin{pmatrix} A^{(1,1)} & * & * & * \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{pmatrix},$$

where N is a $(k-1) \times (k-1)$ matrix. To simplify the notation we write this as

$$PQM = R, \tag{6.28}$$

where the symbols are defined appropriately. By the multiplicative property of determinants we have $D(PQM) = D(P)D(Q)D(M) = (A^{(1,1)})^{k-1}D(M)$ and $D(R) = A^{(1,1)}D(N)$. Hence we have $(A^{(1,1)})^{k-1}D(M) = A^{(1,1)}D(N)$. Take the determinant of both sides of the last equation. Since $|D(N)| = |N|$ by the induction hypothesis, and using $PQM = R$, we find

$$\begin{aligned} |A^{(1,1)}|^{k-1}|D(M)| &= |A^{(1,1)}||D(N)| = |A^{(1,1)}||N| \\ &= |R| = |P||Q||M| = |A^{(1,1)}|^{k-1}|M|. \end{aligned}$$

If $|A^{(1,1)}|$ is neither zero nor a zero divisor, then we can cancel $|A^{(1,1)}|^{k-1}$ from both sides to get Eq. 6.27.

For the general case, we embed K in the polynomial ring$K[z]$, where z is an indeterminate, and replace $A^{(1,1)}$ with the matrix $zI + A^{(1,1)}$. Since the deterinant of $zI+A^{(1,1)}$ is a monic polynomial of degree n, and hence is neither zero nor a zero divisor, Eq. 6.27 holds again. Substituting $z = 0$ (equivalently, equating constant terms of both sides) yields the desired result. □

6.4.2 Tensor Products of Matrices*

Definition: Let $A = (a_{ij}) \in M_{m_1,n_1}(K)$, and let $B \in B_{m_2,n_2}(K)$. Then the *tensor product* or *Kronecker product* of A and B, denoted $A \otimes B \in M_{m_1m_2,n_1n_2}(K)$, is the partitioned matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n_1}B \\ a_{21}B & a_{22}B & \cdots & a_{2n_1}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m_11}B & a_{m_12}B & \cdots & a_{m_1n_1}B \end{pmatrix}. \tag{6.29}$$

It is clear that $I_m \otimes I_n = I_{mn}$.

Lemma Let $A_1 \in M_{m_1,n_1}(K)$, $A_2 \in M_{n_1,r_1}(K)$, $B_1 \in M_{m_2,n_2}(K)$, **and** $B_2 \in M_{n_2,r_2}(K)$. **Then**

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1A_2) \otimes (B_1B_2). \tag{6.30}$$

Proof. Using block multiplication, we see that the (i, j) block of $(A_1A_2) \otimes (B_1B_2)$ is

$$\sum_{k=1}^{n_1} \left((A_1)_{ik}B_1\right)\left((A_2)_{kj}B_2\right) =$$

$$= \left(\sum_{k=1}^{n_1} (A_1)_{ik}(A_2)_{kj} \right) B_1 B_2,$$

which is also seen to be the (i,j) block of $(A_1A_2) \otimes (B_1B_2)$. □

Corollary Let $A \in M_m(K)$ and $B \in M_n(K)$. Then

$$A \otimes B = (A \otimes I_n)(I_m \otimes B). \tag{6.31}$$

and

$$|A \otimes B| = |A|^n |B|^m. \tag{6.32}$$

Proof. Eq. 6.31 is an easy consequence of Eq. 6.30, and then Eq. 6.32 follows easily from the theorem in 6.3.5. □

6.4.3 The Cauchy-Binet Theorem-A Special Version*

The main ingredient in the proof of the Matrix-Tree theorem (see the next section) is the following theorem known as the Cauchy-Binet Theorem. It is more commonly stated and applied with the diagonal matrix $\triangle$ below taken to be the identity matrix. However, the generality given here actually simplifies the proof.

Theorem 6.4.4. *Let A and B be, respectively, $r \times m$ and $m \times r$ matrices, with $r \leq m$. Let $\triangle$ be the $m \times m$ diagonal matrix with entry e_i in the (i,i)-position. For an r-subset S of $[m]$, let A_S and B^S denote, respectively, the $r \times r$ submatrices of A and B consisting of the columns of A, or the rows of B, indexed by the elements of S. Then*

$$det(A \triangle B) = \sum_S det(A_S) det(B^S) \prod_{i \in S} e_i,$$

where the sum is over all r-subsets S of $[m]$.

Proof. We prove the theorem assuming that $e_1, \ldots, e_m$ are independent (commuting) indeterminates over F. Of course it will then hold for all values of $e_1, \ldots, e_m$ in F.

Recall that if $C = (c_{ij})$ is any $r \times r$ matrix over F, then

$$det(C) = \sum_{\sigma \in \mathcal{S}_r} sgn(\sigma) c_{1\sigma(1)} c_{2\sigma(2)} \cdots c_{r\sigma(r)}.$$

Given that $A = (a_{ij})$ and $B = (b_{ij})$, the (i,j)-entry of $A \triangle B$ is $\sum_{k=1}^{m} a_{ik} e_k b_{kj}$, and this is a linear form in the indeterminates $e_1, \ldots, e_m$. Hence $det(A \triangle B)$ is a homogeneous polynomial of degree r in $e_1, \ldots, e_m$. Suppose that $det(A \triangle B)$ has a monomial $e_1^{t_1} e_2^{t_2} \ldots$ where the number of indeterminates e_i that have $t_i > 0$ is less than r. Substitute 0 for the indeterminates e_i that do not appear in $e_1^{t_1} e_2^{t_2} \ldots$, i.e., that have $t_i = 0$. This will not affect the monomial $e_1^{t_1} e_2^{t_2} \ldots$ or its coefficient in $det(A \triangle B)$. But after this substitution $\triangle$ has rank less than r, so $A \triangle B$ has rank less than r, implying that $det(A \triangle B)$ must be the zero polynomial. Hence we see that the coefficient of a monomial in the polynomial $det(A \triangle B)$ is zero unless that monomial is the product of r distinct indeterminates e_i, i.e., unless it is of the form $\prod_{i \in S} e_i$ for some r-subset S of $[m]$.

The coefficient of a monomial $\prod_{i \in S} e_i$ in $det(A \triangle B)$ is found by setting $e_i = 1$ for $i \in S$, and $e_i = 0$ for $i \notin S$. When this substitution is made in $\triangle$, $A \triangle B$ evaluates to $A_S B^S$. So the coefficient of $\prod_{i \in S} e_i$ in $det(A \triangle B)$ is $det(A_S) det(B^S)$. □

Exercise 6.4.4.1. *Let M be an $n \times n$ matrix all of whose linesums are zero. Then one of the eigenvalues of M is $\lambda_1 = 0$. Let $\lambda_2, \ldots, \lambda_n$ be the other eigenvalues of M. Show that all principal $n-1$ by $n-1$ submatrices have the same determinant and that this value is $\frac{1}{n}\lambda_2 \lambda_3 \cdots \lambda_n$.*

Sketch of Proof: First note that since all line sums are equal to zero, the entries of the matrix are completely determined by the entries of M in the first n rows and first n columns, and that the entry in the (n, n) position is the sum of all $(n-1)^2$ entries in the first $n-1$ rows and columns.

Clearly $\lambda_1 = 0$ is an eigenvalue of A. Observe the appearance of the $(n-1) \times (n-1)$ subdeterminant obtained by deleting the bottom row and right hand column. Then consider the principal subdeterminant obtained by deleting row j and column j, $1 \leq j \leq n-1$, from the original matrix M. In this $(n-1) \times (n-1)$ matrix, add the first $n-2$ columns to the last one, and then add the first $n-2$ rows to the last one. Now multiply the last column and the last row by -1. This leaves a matrix that could have been obtained from the original upper $(n-1) \times (n-1)$ submatrix by moving its jth row and column to the last positions. So it has the same determinant.

Now note that the coefficient of x in the characteristic polynomial $f(x) = det(xI - A)$ is $(-1)^{n-1} \lambda_2 \lambda_3 \cdots \lambda_n$, since $\lambda_1 = 0$, and it is also $(-1)^{n-1} \sum det(B)$, where the sum is over all principal subdeterminants of order $n-1$, which by

the previous paragraph all have the same value. Hence $det(B) = \frac{1}{n}\lambda_2\lambda_3 \cdots \lambda_n$, for any principal subdeterminant of order $n-1$.

6.4.5 The Matrix-Tree Theorem*

The "matrix-tree" theorem expresses the number of spanning trees in a graph as the determinant of an appropriate matrix, from which we obtain one more proof of Cayley's theorem counting labeled trees.

An **incidence matrix** N of a directed graph H is a matrix whose rows are indexed by the vertices V of H, whose columns are indexed by the edges E of H, and whose entries are defined by:

$$N(x,e) = \begin{cases} 0 & \text{if } x \text{ is not incident with } e, \text{ or e is a loop,} \\ 1 & \text{if } x \text{ is the head of } e, \\ -1 & \text{if } x \text{ is the tail of } e. \end{cases}$$

Lemma 6.4.6. *If H has k components, then $rank(N) = |V| - k$.*

Proof. N has $v = |V|$ rows. The rank of N is $v - n$, where n is the dimension of the left null space of N, i.e., the dimension of the space of row vectors g for which $gN = 0$. But if e is any edge, directed from x to y, then $gN = 0$ if and only if $g(x) - g(y) = 0$. Hence $gN = 0$ iff g is constant on each component of H, which says that n is the number k of components of H. □

Lemma 6.4.7. *Let A be a square matrix that has at most two nonzero entries in each column, at most one 1 in each column, at most one -1 in each column, and whose entries are all either 0, 1 or -1. Then $det(A)$ is 0, 1 or -1.*

Proof. This follows by induction on the number of rows. If every column has both a 1 and a -1, then the sum of all the rows is zero, so the matrix is singular and $det(A) = 0$. Otherwise, expand the determinant by a column with one nonzero entry, to find that it is equal to ± 1 times the determinant of a smaller matrix with the same property. □

Corollary 6.4.8. *Every square submatrix of an incidence matrix of a directed graph has determinant 0 or ± 1. (Such a matrix is called* **totally unimodular**.*)*

Theorem 6.4.9. *(The* Matrix-Tree Theorem*) The number of spanning trees in a connected graph G on n vertices and without loops is the determinant of any $n-1$ by $n-1$ principal submatrix of the matrix $D-A$, where A is the adjacency matrix of G and D is the diagonal matrix whose diagonal contains the degrees of the corresponding vertices of G.*

Proof. First let H be a connected digraph with n vertices and with incidence matrix N. H must have at least $n-1$ edges, because it is connected and must have a spanning tree, so we may let S be a set of $n-1$ edges. Using the notation of the Cauchy-Binet Theorem, consider the n by $n-1$ submatrix N_S of N whose columns are indexed by elements of S. By Lemma 6.4.6, N_S has rank $n-1$ iff the spanning subgraph of H with S as edge set is connected, i.e., iff S is the edge set of a tree in H. Let N' be obtained by dropping any single row of the incidence matrix N. Since the sum of all rows of N (or of N_S) is zero, the rank of N'_S is the same as the rank of N_S. Hence we have the following:

$$det(N'_S) = \begin{cases} \pm 1 & \text{if } S \text{ is the edge set of a spanning tree in } H, \\ 0 & \text{otherwise.} \end{cases} \tag{6.33}$$

Now let G be a connected loopless graph on n vertices. Let H be any digraph obtained by orienting G, and let N be an incidence matrix of H. Then we claim $NN^T = D - A$. For,

$$\begin{aligned} (NN^T)_{xy} &= \textstyle\sum_{e \in E(G)} N(x,e)N(y,e) \\ &= \begin{cases} \deg(x) \text{ if } x = y, \\ -t \qquad\qquad\quad \text{if } x \text{ and } y \text{ are joined by } t \text{ edges in } G. \end{cases} \end{aligned}$$

An $n-1$ by $n-1$ principal submatrix of $D-A$ is of the form $N'N'^T$ where N' is obtained from N by dropping any one row. By Cauchy-Binet,

$$det(N'N'^T) = \sum_S det(N'_S)det(N'^T_S) = \sum_S (det(N'_S))^2,$$

where the sum is over all $n-1$ subsets S of the edge set. By Eq. 6.33 this is the number of spanning trees of G. □

Exercise 6.4.9.1. *(Cayley's Theorem) In the Matrix-Tree Theorem, take G to be the complete graph K_n. Here the matrix $D-A$ is $nI-J$, where I is the identity matrix of order n, and J is the n by n matrix of all 1's. Now calculate the determinant of any $n-1$ by $n-1$ principal submatrix of this matrix to obtain another proof that K_n has n^{n-2} spanning trees.*

Exercise 6.4.9.2. *In the statement of the Matrix-Tree Theorem it is not necessary to use* principal *subdeterminants. If the $n-1 \times n-1$ submatrix M is obtained by deleting the ith row and jth column from $D-A$, then the number of spanning trees is $(-1)^{i+j}det(M)$. This follows from the more general lemma: If A is an $n-1 \times n$ matrix whose row sums are all equal to 0 and if A_j is obtained by deleting the jth column of A, $1 \leq j \leq n$, then $det(A_j) = -det(A_{j+1})$.*

6.4.10 The Cauchy-Binet Theorem - A General Version*

Let $1 \leq p \leq m \in \mathcal{Z}$, and let $Q_{p,m}$ denote the set of all sequences $\alpha = (i_1, i_2, \cdots, j_p)$ of p integers with $1 \leq i_1 < i_2 < \cdots < i_p \leq m$. Note that $|Q_{p,m}| = \binom{m}{p}$.

Let K be a commutative ring with 1, and let $A \in M_{m,n}(K)$. If $\alpha \in Q_{p,m}$ and $\beta \in Q_{j,n}$, let $A[\alpha|\beta]$ denote denote the submatrix of A consisting of the element whose row index is in α and whose column index is in β. If $\alpha \in Q_{p,m}$, then there is a complementary sequence $\hat{\alpha} \in Q_{m-p,m}$ consisting of the list of exactly those positive integers between 1 and m that are not in α, and the list is in increasing order.

Theorem 6.4.11. *Let $A \in M_{m,n}(K)$ and $B \in M_{n,p}(K)$. Assume that $1 \leq t \leq min\{m,n,p\}$ and let $\alpha \in Q_{t,m}$, $\beta \in Q_{t,p}$. Then*

$$det(AB[\alpha|\beta]) = \sum_{\gamma \in Q_{t,n}} det(A[\alpha|\gamma]) \cdot det(B[\gamma|\beta]).$$

Proof. Suppose that $\alpha = (\alpha_1, \ldots, \alpha_t)$, $\beta = (\beta_1, \ldots, \beta_t)$, and let $C = AB[\alpha|\beta]$. Then

$$C_{ij} = \sum_{k=1}^{n} a_{\alpha_i k} b_{k \beta_j}.$$

So we have

$$C = \begin{pmatrix} \sum_{k=1}^{n} a_{\alpha_1 k} b_{k\beta_1} & \cdots & \sum_{k=1}^{n} b_{k\beta_t} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{\alpha_t k} b_{k\beta_1} & \cdots & \sum_{k=1}^{n} a_{\alpha_t k} b_{k\beta_t} \end{pmatrix}.$$

To calculate the determinant of C we start by using the n-linearity in the first row, then the second, row, etc.

$$det(C) = \sum_{k_1=1}^{n} a_{\alpha_1 k_1} \cdot det \begin{pmatrix} b_{k_1\beta_1} & \cdots & b_{k_1\beta_t} \\ \sum_{k=1}^{n} a_{\alpha_2 k} b_{k\beta_1} & \cdots & \sum_{k=1}^{n} a_{\alpha_2 k} b_{k\beta_t} \\ \vdots & \cdots & \vdots \sum_{k=1}^{n} a_{\alpha_2 k} b_{k\beta_t} \\ \sum_{k=1}^{n} a_{\alpha_t k} b_{k\beta_1} & \cdots & \sum_{k=1}^{n} a_{\alpha_t k} b_{k\beta_t} \end{pmatrix} =$$

$$= \sum_{k_1=1}^{n} \cdots \sum_{k_t=1}^{n} a_{\alpha_1 k_1} \cdots a_{\alpha_t k_t} \cdot det \begin{pmatrix} b_{k_1\beta_1} & \cdots & b_{k_1\beta_t} \\ \vdots & & \vdots \\ b_{k_t\beta_1} & \cdots & b_{k_t\beta_t} \end{pmatrix}. \tag{6.34}$$

If $k_i = k_j$ for $i \neq j$, then

$$det \begin{pmatrix} b_{k_1\beta_1} & \cdots & b_{k_1\beta_t} \\ \vdots & & \vdots \\ b_{k_t\beta_1} & \cdots & b_{k_t\beta_t} \end{pmatrix} = 0.$$

The the only possible nonzero determinant occurs when the $(k_1, \ldots, k_t)$ is a permutation of a sequence $\gamma = (\gamma_1, \ldots, \gamma_t) \in Q_{t,n}$. Let $\sigma \in \mathcal{S}_t$ be the permutation of $\{1, 2, \ldots, t\}$ such that $\gamma_i = k_{\sigma(i)}$ for $1 \leq i \leq t$. Then

$$det \begin{pmatrix} b_{k_1\beta_1} & \cdots & b_{k_1\beta_t} \\ \vdots & & \vdots \\ b_{k_t\beta_1} & \cdots & b_{k_t\beta_t} \end{pmatrix} sgn(\sigma) det(B[\gamma|\beta]). \tag{6.35}$$

Given a fixed $\gamma \in Q_{t,n}$, all possible permutations of γ are included in the summation in Eq. 6.34. Therefore Eq. 6.34 may be rewritten , using Eq. 6.35, as

$$det(C) = \sum_{\gamma \in Q_{t,n}} \left(\sum_{\sigma \in \mathcal{S}_t} sgn(\sigma) a_{\alpha_1 \gamma_{\sigma(1)}} \cdots a_{\alpha_t \gamma_{\sigma(t)}} \right) det(B[\gamma|\beta]),$$

which is the desired formula. □

The Cauchy-Binet formula gives another verification of the fact that $det(AB) = det(A) \cdot det(B)$ for square matrices A and B.

6.4.12 The General Laplace Expansion*

For $\gamma = (\gamma_1, \ldots, \gamma_t) \in Q_{t,n}$, put $s(\gamma) = \sum_{j=1}^{t} \gamma_j$.

Theorem 6.4.13. *Let $A \in M_n(K)$ and let $\alpha \in Q_{t,n}$ $(1 \leq t \leq n)$ be given. Then*

$$det(A) = \sum_{\gamma \in Q_{t,n}} (-1)^{s(\alpha)+s(\gamma)} det(A[\alpha|\gamma]) \cdot det(A[\hat{\alpha}|\hat{\gamma}]). \tag{6.36}$$

Proof. For $A \in M_n(K)$, define

$$D_\alpha(A) = \sum_{\gamma \in Q_{t,n}} (-1)^{s(\alpha)+s(\gamma)} det(A[\alpha|\gamma]) \cdot det(A[\hat{\alpha}|\hat{\gamma}]). \tag{6.37}$$

Then $D_\alpha : M_n(K) \to K$ is easily shown to be n-linear as a function on the *columns* of A. To complete the proof, it is only necessary to show that D_α is alternating and that $D_\alpha(I_n) = 1$. Thus, suppose that the columns of A labeled p and q, $p < q$, are equal. If p and q are both in $\gamma \in Q_{t,n}$, then $A[\alpha|\gamma]$ will have two columns equal and hence have zero determinant. Similarly, if p and q are both in $\hat{\gamma} \in Q_{n-t,n}$, then $det(A[\hat{\alpha}|\hat{\gamma}) = 0$. Thus in the evaluation of $D_\alpha(A)$ it is only necessary to consider those $\gamma \in Q_{t,n}$ such that $p \in \gamma$ and $q \in \hat{\gamma}$, or vice versa. So suppose $p \in \gamma$, $q \in \hat{\gamma}$, and defind a new sequence γ' in $Q_{t,n}$ by replacing $p \in \gamma$ by q. Thus $\hat{\gamma}'$ agrees with $\hat{\gamma}$ except that q has been replaced by p. Thus

$$s(\gamma') - s(\gamma) = q - p. \tag{6.38}$$

(Note that $s(\gamma) - p$ and $s(\gamma') - q$ are both the sum of all the things in γ except for p.) Now consider the sum

$$(-1)^{s(\gamma)} det(A[\alpha|\gamma]) det(A[\hat{\alpha}|\hat{\gamma}]) + (-1)^{s(\gamma')} det(A[\alpha|\gamma']) det(A[\hat{\alpha}|\hat{\gamma}']),$$

which we denote by $S(A)$. We claim that this sum is 0. Assuming this, since γ and γ' appear in pairs in $Q_{t,n}$, it follows that $D_\alpha(A) = 0$ whenever two columns of A agree, forcing D_α to be alternating. So now we show that $S(A) = 0$.

Suppose that $p = \gamma_k$ and $q = \hat{\gamma}_l$. Then γ and γ' agree except in the range from p to q, as do $\hat{\gamma}$ and $\hat{\gamma}'$. This includes a total of $q - p + 1$ entries. If r of these entries are included in γ, then

$$\gamma_1 < \cdots < \gamma_k = p < \gamma_{k+1} < \cdots < \gamma_{k+r-1} < q < \gamma_{k+r} < \cdots < \gamma_t$$

and

$$A[\alpha|\gamma'] = A[\alpha|\gamma]P_{w^{-1}},$$

where w is the r-cycle $(k+r-1, k+r-2, \ldots, k)$. Similarly,

$$A[\hat{\alpha}|\hat{\gamma}'] = A[\hat{\alpha}|\hat{\gamma}]P_{w'}$$

where w' is a $(q-p+1-r)$-cycle. Thus,

$$(-1)^{s(\gamma')} det(A[\alpha|\gamma'])det(A[\hat{\alpha}|\hat{\gamma}']) =$$
$$= (-1)^{s(\gamma')+(r-1)+(q-p)-r} det(A[\alpha|\gamma])det(A[\hat{\alpha}|\hat{\gamma}]).$$

Since $s(\gamma') + (q-p) - 1 - s(\gamma) = 2(q-p) - 1$ is odd, we conclude that $S(A) = 0$. Thus D_α is n-linear and alternating. It is routine to check that $D_\alpha(I_n) = 1$, completing the proof. □

Applying this formula for $det(A)$ to $det(A^T)$ gives the Laplace expansion in terms of columns.

6.4.14 Determinants, Ranks and Linear Equations*

If K is a commutative ring with 1 and $A \in M_{m,n}(K)$, and if $1 \leq t \leq \min\{m,n\}$, then a $t \times t$ *minor* of A is the determinant of any submatrix $A[\alpha|\beta]$ where $\alpha \in Q_{t,m}$, $\beta \in Q_{t,n}$. The *determinantal rank* of A, denoted D-rank(A), is the largest t such that there is a nonzero $t \times t$ minor of A.

With the same notation,

$$F_t(A) = \langle \{detA[\alpha|\beta] : \alpha \in Q_{t,m}, \beta \in Q_{t,n}\} \rangle \subseteq K.$$

That is, $F_t(A)$ is the ideal of K generated by all the $t \times t$ minors of A. Put $F_0(A) = K$ and $F_t(A) = 0$ if $t > \min\{m,n\}$. $F_t(A)$ is called the t^{th}-Fitting ideal of A. The Laplace expansion of determinants along a row or column shows that $F_{t+1}(A) \subseteq F_t(A)$. Thus there is a decreasing chain of ideals

$$K = F_0(A) \supseteq F_1(A) \supseteq F_2(A) \supseteq \cdots.$$

Definition: If K is a PID, then $F_t(A)$ is a principal ideal, say $F_t(A) = \langle d_t(A) \rangle$ where $d_t(A)$ is the greatest common divisor of all the $t \times t$ minors of

A. In this case, a generator of $F_t(A)$ is called the t^{th}-*determinantal divisor of* A.

Definition If $A \in M_{m,n}(K)$, then the M-rank(A) is defined to be the largest t such that $\{0\} = Ann(F_t(A)) = \{k \in K : kd = 0 \text{ for all } d \in F_t(A)\}$.

Obs. 6.4.15. *1. M-rank$(A) = 0$ means that $Ann(F_1(a)) \neq \{0\}$. That is, there is a nonzero $a \in K$ with $a \cdot a_{ij} = 0$ for all entries a_{ij} of A. Note that this is stronger than saying that every element of A is a zero divisor. For example, if $A = (2\ 3) \in M_{1,2}(\mathcal{Z}_6)$, then every element of A is a zero divisor in $\mathcal{Z}_6$, but there is no single nonzero element of $\mathcal{Z}_6$ that annihilates both entries in the matrix.*

2. If $A \in M_n(K)$, then M-rank$(A) = n$ means that $det(A)$ is not a zero divisor of K.

3. To say that M-rank$(A) = t$ means that there is an $a \neq 0 \in K$ with $a \cdot D = 0$ for all $(t+1) \times (t+1)$ minors D of A, but there is no nonzero $b \in K$ which annihilates all $t \times t$ minors of A by multiplication. In particular, if $det(A[\alpha|\beta])$ is not a zero divisor of K for some $\alpha \in Q_{s,m}$, $\beta \in Q_{s,n}$, then M-rank$(A) \geq s$.

Lemma 6.4.16. *If $A \in M_{m,n}(K)$, then*

$$0 \leq M - rank(A) \leq D - rank(A) \leq min\{m, n\}.$$

Proof. Routine exercise. □

We can now give a criterion for solvability of the homogeneous linear equation $AX = \overline{0}$, where $A \in M_{m,n}(K)$. This equation always has the trivial solution $X = \overline{0}$, so we want a criterion for the existence of a solution $X \neq \overline{0} \in M_{n,1}(K)$.

Theorem 6.4.17. *Let K be a commutative ring with 1 and let $A \in M_{m,n}(K)$. The matrix equation $AX = \overline{0}$ has a nontrivial solution $X \neq \overline{0} \in M_{n,1}(K)$ if and only if*

$$M - rank(A) < n.$$

Proof. Suppose that M-rank$(A) = t < n$. then Ann$(F_{t+1}(A)) \neq \{0\}$, so choose $b \neq 0 \in K$ with $b \cdot F_{t+1}(A) = \{0\}$. Without loss of generality, we may assume that $t < m$, since, if necessary, we may replace the system $AX = \overline{0}$

with an equivalent one (i.e., one with the same solutions) by adding some rows of zeroes to the bottom of A. If $t = 0$, then $ba_{ij} = 0$ for all a_{ij} and we may take

$$X = \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix}.$$

Then $X \neq \overline{0} \in M_{n,1}(K)$ and $AX = \overline{0}$.

So suppose that $t > 0$. Then $b \notin Ann(F_t(A)) = \{0\}$, so $b \cdot det(A[\zeta|\beta]) \neq 0$ for some $\alpha \in Q_{t,m}$, $\beta \in Q_{t,n}$. By permuting rows and columns, which does note affect whether $AX = \overline{0}$ has a nontrivial solution, we can assume $\alpha = (1, \ldots, t) = \beta$. For $1 \leq i \leq t+1$ let $\beta_i = (1, 2, \ldots, \hat{i}, \ldots, t+1) \in Q_{t,t+1}$, where $\hat{i}$ indicates that i is deleted. Let $d_i = (-1)^{t+1+i} det(A[\alpha|\beta_i])$. Thus $d_1, \ldots, d_{t+1}$ are the cofactors of the matrix

$$A_1 = A[(1, \ldots, t+1)|(1, \ldots, t+1)]$$

obtained by deleting row $t+1$ and column i. Hence the Laplace expansion gives

$$\begin{cases} \sum_{j=1}^{t+1} a_{ij} d_j = 0, & \text{if } 1 \leq i \leq t, \\ \sum_{j=i}^{t+1} a_{ij} d_j = \det(A[(1, \ldots, t, i)|(1, \ldots, t, t+1)]), & \text{if } t < i \leq m. \end{cases} \tag{6.39}$$

Let $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, where

$$\begin{cases} x_i = bd_i, & \text{if } 1 \leq i \leq t+1, \\ x_i = 0, & \text{if } t+2 \leq i \leq n. \end{cases}$$

Then $X \neq \overline{0}$ since $x_{t+1} = b \cdot \det(A[\alpha|\beta]) \neq 0$. But Eq. 6.39 and the fact that $b \in Ann(F_{t+1}(A))$ show that

$$AX = \begin{pmatrix} b \sum_{j=1}^{t+1} a_{1j} d_j \\ \vdots \\ b \sum_{j=1}^{t+1} a_{mj} d_j \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \det(A[(1,\ldots,t,t+1)|(1,\ldots,t,t+1)]) \\ \vdots \\ b \det(A[(1,\ldots,t,m)|(1,\ldots,t,t+1)]) \end{pmatrix} = 0.$$

Thus X is a nontrivial solution to the equation $AX = \overline{0}$.

Conversely, assume that $X \neq \overline{0} \in M_{n,1}(K)$ is a nontrivial solution to $AX = \overline{0}$, and choose k with $x_k \neq 0$. We claim that $\text{Ann}(F_n(A)) \neq \{0\}$. If $n > m$, then $F_n(A) = \{0\}$, and hence $\text{Ann}(F_n(A)) = K \neq (0)$. Thus we may assume that $n \leq m$. Let $\alpha = (1,\ldots,n)$ and for each $\beta \in Q_{n,m}$, let $B_\beta = A[\alpha|\beta]$. Then since $AX = \overline{0}$ and since each row of B_β is a full row of A, we conclude that $B_\beta = \overline{0}$. The adjoint matrix formula (Eq. 6.20) then shows that

$$(det(B_\beta))X = (Adj\ B_\beta)B_\beta X = \overline{0},$$

from which we conclude that $x_k\ det(B_\beta) = 0$. Since $\beta \in Q_{n,m}$ is arbitrary, we conclude that $x_k \cdot F_n(A) = 0$, i.e., $x_k \in \text{Ann}(F_n(A))$. But $x_k \neq 0$, so $\text{Ann}(F_n(A)) \neq \{0\}$, and we conclude that $M\text{-rank}(A) < n$, completing the proof. □

In case K is an integral domain we may replace the M-rank by the ordinary determinantal rank to conclude the following:

Corollary 6.4.18. *If K is an integral domain and $A \in M_{m,n}(K)$, then $AX = \overline{0}$ has a nontrivial solution if and only if $D\text{-rank}(A) < n$.*

Proof. If $I \subseteq K$, then $\text{Ann}(I) \neq \{0\}$ if and only if $I = \{0\}$ since an integral domain has no nonzero zero divisors. Therefore, in an integral domain $D\text{-rank}(A) = M\text{-rank}(A)$. □

The results for n equations are in n unknowns are even simpler.

Corollary 6.4.19. *Let K be a commutative ring with 1.*

1. *If $A \in M_n(K)$, then $AX = \overline{0}$ has a nontrivial solution if and only if $det(A)$ is a zero divisor of K.*

2. *If K is an integral domain and $A \in M_n(K)$, then $AX = \overline{0}$ has a nontrivial solution if and only if $det\ (A) = 0$.*

Proof. If $A \in M_n(K)$, then $F_n(A) = \langle det(A) \rangle$, so M-rank$(A) < n$ if and only if $det(A)$ is a zero divisor. In particular, if K is an integral domain then M-rank$(A) < n$ if and only if $det(A) = 0$. □

There are still two other concepts of rank which can be defined for matrices with entries in a commutative ring.

Definition Let K be a commutative ring with 1 and let $A \in M_{m,n}(K)$. Then we will define the *row rank* of A, denoted by *row-rank*(A), to be the maximum number of linearly independent rows in A, while the *column rank* of A, denoted *col-rank*(A), is the maximum number of linearly independent columns.

Corollary 6.4.20. *1. If K is a commutative ring with 1 and $A \in A_{m,n}(K)$, then*

$$max\{row\text{-}rank(A), col\text{-}rank(A) \leq M - rank(A) \leq D - rank(A).$$

2. If K is an integral domain, then

$$row\text{-}rank(A) = col\text{-}rank(A) = M - rank(A) = D - rank(A).$$

Proof. We sketch the proof of the result when K is an integral domain. In fact, it is possible to embed K in its field F of quotients and do all the algebra in F. Recall the following from an undergraduate linear algebra course. The proofs work over any field, even if you did only consider them over the real numbers. Let A be an $m \times n$ matrix with entries in F. Row reduce A until arriving at a matrix R in row-reduced echelon form. The first thing to remember here is that the row space of A and the row space of R are the same. Similarly, the right null space of A and the right null space of R are the same. (Warning: the column space of A and that of R usually are not the same!) So the leading (i.e., leftmost) nonzero entry in each nonzero row of R is a 1, called a *leading* 1. Any column with a leading 1 has that 1 as its only nonzero entry. The nonzero rows of R form a basis for the row space of A, so the number r of them is the row-rank of A. The (right) null space of R (and hence of A) has a basis of size $n - r$. Also, one basis of the column space of A is obtained by taking the set of columns of A in the positions now indicated by the columns of R in which there are leading 1's. So the column rank of A is also r. This has the interesting corollary that if any r independent columns of A are selected, there must be some

r rows of those columns that are linearly independent, so there is an $r \times r$ submatrix with rank r. Hence this submatrix has determinant different from 0. Conversely, if some $r \times r$ submatrix has determinant different from 0, then the "short" columns of the submatrix must be independent, so the "long" columns of A to which they belong must also be independent. It is now clear that the row-rank, column-rank, M-rank and determinantal rank of A are all the same. □

Obs. 6.4.21. *Since all four ranks of A are the same when K is an integral domain, in this case we may speak unambiguously of the* rank *of A, denoted rank(A). Moreover, the condition that K be an integral domain is truly necessary, as the following example shows.*

Define $A \in M_4(\mathcal{Z}_{210})$ by

$$A = \begin{pmatrix} 0 & 2 & 3 & 5 \\ 2 & 0 & 6 & 0 \\ 3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}.$$

It is an interesting exercise to show the following:

1. *row-rank(A) = 1.*
2. *col-rank(A) = 2.*
3. *M-rank(A) = 3.*
4. *D-rank(A) = 4.*

Theorem 6.4.22. *Let K be a commutative ring with 1, let M be a finitely generated K-module, and let $S \subseteq M$ be a subset. If $|S| > \mu(M) = rank(M)$, then S is K-linearly dependent.*

Proof. Let $\mu(M) = m$ and let $T = \{w_1, \ldots, w_m\}$ be a generating set for M consisting of m elements. Choose n distinct elements $\{v_1, \ldots, v_n\}$ of S for some $n > m$, which is possible by hypothesis. Since $M = \langle w_1, \ldots, w_m \rangle$, we may write

$$v_j = \sum_{i=1}^{m} a_{ij} w_i, \text{ with } a_{ij} \in K.$$

Let $A = (a_{ij}) \in M_{m,n}(K)$. Since $n > m$, it follows that M-rank$(A) \leq m < n$, so Theorem 6.4.17 shows that there is an $X \neq \overline{0} \in M_{n,1}(K)$ such that $AX = \overline{0}$. Then

$$\begin{aligned} \sum_{j=1}^{n} x_j v_j &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} w_i \right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) w_i \qquad (6.40) \\ &= 0, \text{ since } AX = \overline{0}. \end{aligned}$$

Therefore, S is K-linearly dependent. □

Corollary 6.4.23. *Let K be a commutative ring with 1, let M be a K-module, and let $N \subseteq M$ be a free submodule. Then rank$(N) \leq$ rank(M).*

Proof. If rank$(M) = \infty$, there is nothing to prove, so assume that rank(M) $= m < \infty$. If rank$(N) > m$, then there is a linearly independent subset of M, namely a basis of N, with more than m elements, which contradicts Theorem 6.4.22. □

Theorem 6.4.24. *If $A \in M_{m,n}(K)$ and $B \in M_{n,p}(K)$, then*

$$D - rank(AB) \leq min\{D - rank(A), D - rank(B)\}. \qquad (6.41)$$

Proof. Let $t > \min\{D - \text{rank}(A), D - \text{rank}(B)\}$ and suppose that $\alpha \in Q_{t,m}$, $\beta \in Q_{t,p}$. Then by the Cauchy-Binet formula

$$det(AB[\alpha|\beta]) = \sum_{\gamma \in Q_{t,n}} det(A[\alpha|\gamma]) det(B[\gamma|\beta]).$$

Since $t > \min\{D - \text{rank}(A), D - \text{rank}(B)\}$, at least one of the determinants $det(A[\alpha|\beta])$ or $det(B[\gamma|\beta])$ must be 0 for each $\gamma \in Q_{t,n}$. Thus $det(AB[\alpha|\beta]) = 0$, and since α and β are arbitrary, it follows that D-rank$(AB) < t$, as required.

□

The preceding theorem has a useful corollary.

Corollary 6.4.25. *Let $A \in M_{m,n}(K)$, $U \in GL(m,K)$, $V \in GL(n,K)$. Then*

$$D - \mathit{rank}(UAV) = D - \mathit{rank}(A).$$

Proof. Any matrix $B \in M_{m,n}(K)$ satisfies D-rank$(B) \leq \min\{m,n\}$. Since D-rank$(U) = m$ and D-rank$(V) = n$, it follows from Eq. 6.41 that

$$D - \operatorname{rank}(UAV) \leq \min\{D - \operatorname{rank}(A), n, m\} = D - \operatorname{rank}(A)$$

and

$$D - \operatorname{rank}(A) = D - \operatorname{rank}(U^{-1}(UAV)V^{-1}) \leq D - \operatorname{rank}(UAV).$$

This completes the proof. □

6.5 Exercises

1. (Vandermonde Determinant)

   Let $t_1, \ldots, t_n$ be commuting indeterminates over K, and let A be the $n \times n$ matrix whose entries are from the commutative ring $K[t_1, \ldots, t_n]$ defined by

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-1} \end{pmatrix}.$$

   Then

$$\det A = \prod_{1 \leq j < i \leq n} (t_i - t_j).$$

Chapter 7

Operators and Invariant Subspaces

Throughout this chapter F will be an arbitrary field unless otherwise restricted, and V will denote an arbitrary vectors space over F. Much of our work will require that V be finite dimensional, but we shall be general as long as possible. Given an operator $T \in \mathcal{L}(V)$, our main interest will be in finding T-invariant subspaces $U_1, \ldots, U_r$ such that $V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$. Since each U_i is invariant under T, we may consider the restriction $T_i = T|_{U_i}$ of T to the subspace U_i. Then we write $T = T_1 \oplus \cdots \oplus T_r$. It is usually easier to analyze T by writing it as the sum of the operators T_i and then analyzing the operators T_i on the smaller subspaces U_i.

7.1 Eigenvalues and Eigenvectors

Let $T \in \mathcal{L}(V)$. Then $\{\vec{0}\}$, V, null(T), and Im(T) are invariant under T. However, often these are not especially interesting as invariant subspaces, and we want to begin with 1-dimensional T-invariant subspaces.

Suppose that U is a 1-dimensional T-invariant subspace. Then there exists some nonzero vector $u \in U$. Since $T(u) \in U$, there is some scalar $a \in F$ such that $T(u) = au$. By hypothesis (u) is a basis for U. If bu is any vector in U, then $T(bu) = bT(u) = b(au) = a(bu)$. Hence for each $v \in U$, $T(v) = av$. If $a \in F$ satisfies the property that there is some nonzero vector $v \in V$ such that $T(v) = av$, then a is called an *eigenvalue* of T. A vector $v \in V$ is called an *eigenvector of* T belonging to the eigenvalue λ provided

$T(v) = \lambda v$, i.e., $(T - \lambda I)(v) = \vec{0}$. Note that the eigenvalue λ could be the zero scalar, but it is an eigenvalue if and only if there is a nonzero eigenvector belonging to it. For this reason many authors restrict all eigenvectors to be nonzero. However, it is convenient to include $\vec{0}$ in the set of eigenvectors belonging to any particular eigenvalue so that the set of all eigenvectors belonging to some eigenvalue is a subspace, in fact a T-invariant subspace.

We have the following:

Obs. 7.1.1. *If λ is an eigenvalue of T, then null$(T - \lambda I)$ is the T-invariant subspace of V consisting of all eigenvectors belonging to λ.*

Consider the example $T \in \mathcal{L}(F^2)$ defined by $T(y, z) = (2z, -y)$. Then $T(y, z) = \lambda(y, z)$ if and only if $2z = \lambda y$ and $-y = \lambda z$, so $2z = \lambda(-\lambda z)$. If $(y, z) \neq (0, 0)$, then $\lambda^2 = -2$. If $F = \mathcal{R}$, for example, then T has no eigenvalue. However, if F is algebraically closed, for example, then T has two eigenvalues $\pm\lambda$ where λ is one of the two solutions to $\lambda^2 = -2$. If $F = \mathcal{Z}_5$, then $-2 = 3$ is a non-square in F, so T has no eigenvalues. But if $F = \mathcal{Z}_{11}$, then $\lambda = \pm 3$ satisfies $\lambda^2 = -2$ in F.

Nonzero eigenvectors belonging to distinct eigenvalues are linearly independent.

Theorem 7.1.2. *Let $T \in \mathcal{L}(V)$ and suppose that $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T with corresponding nonzero eigenvectors $v_1, \ldots, v_m$. Then the list $(v_1, \ldots, v_m)$ is linearly independent.*

Proof. Suppose $(v_1, \ldots, v_m)$ is linearly dependent. By the Linear Dependence Lemma we may let j be the smallest positive integer for which $(v_1, \ldots, v_j)$ is linearly dependent, so that $v_j \in \text{span}(v_1, \ldots, v_{k-1})$. So there are scalars $a_1, \ldots, a_{j-1}$ for which

$$v_j = a_1 v_1 + \cdots + a_{j-1} v_{j-1}. \tag{7.1}$$

Apply T to both sides of this equation to obtain

$$\lambda_j v_j = a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_{j-1} \lambda_{j-1} v_{j-1}.$$

Multiply both sides of Eq. 7.1 by λ_j and subtract the equation above from it. This gives

$$\vec{0} = a_1(\lambda_j - \lambda_1)v_1 + \cdots + a_{j-1}(\lambda_j - \lambda_{j-1})v_{j-1}.$$

Because j was chosen to be the smallest integer for which $v_j \in \text{span}(v_1, \ldots, v_{j-1})$ we now have that $(v_1, \ldots, v_{j-1})$ is linearly independent. Since the λ's were all distinct, this means that $a_1 = \cdots = a_{j-1} = 0$, implying that $v_j = \vec{0}$ (by Eq. 7.1), contradicting our hypothesis that all the v_i's are nonzero. Hence our assumption that $(v_1, \ldots, v_j)$ is linearly dependent must be false. □

Since each linearly independent set of a finite dimensional space has no more elements than the dimension of that space, we obtain the following result.

Corollary 7.1.3. *If $dim(V) = n < \infty$, then V can never have more than n distinct eigenvalues.*

7.2 Upper-Triangular Matrices

In Chapter 5 we applied polynomials to elements of some linear algebra over F. In particular, if $T \in \mathcal{L}(V)$ and $f, g, h \in F[x]$ with $f = gh$, then by Theorem 5.2.5 we know that $f(T) = g(T)h(T)$.

Theorem 7.2.1. *Let V be a nonzero, finite dimensional vector space over the algebraically closed field F. Then each operator T on V has an eigenvalue.*

Proof. Suppose $\dim(V) = n > 0$ and choose a nonzero vector $v \in V$. Let $T \in \mathcal{L}(V)$. Then the set

$$(v, T(v), T^2(v), \ldots, T^n(v))$$

of $n+1$ vectors in an n-dimensional space cannot be linearly independent. So there must be scalars, not all zero, such that $\vec{0} = a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_n T^n(v)$. Let m be the largest index such that $a_m \neq 0$. Since $v \neq \vec{0}$, the coefficients $a_1, \ldots, a_n$ cannot all be 0, so $0 < m \leq n$. Use the $a's$ to construct a polynomial which can be written in factored form as

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m = c(z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m),$$

where $c \in F$ is nonzero, each $\lambda_j \in F$, and the equation holds for all $z \in F$. We then have

$$\begin{aligned} \vec{0} &= a_0 v + a_1 T(v) + \cdots + a_m T^m(v) \\ &= (a_0 I + a_1 T + \cdots a_m T^m)(v) \\ &= c(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_m I)(v), \end{aligned} \tag{7.2}$$

which means that $T - \lambda_j$ is not injective for at least one j, i.e., T has at least one eigenvalue. □

Theorem 7.2.2. *Suppose $T \in \mathcal{L}(V)$ and $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of V. Then the following are equivalent:*

(i) $[T]_\mathcal{B}$ is upper triangular.

(ii) $T(v_k) \in span(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$.

(iii) The $span(v_1, \ldots, v_k)$ is T-invariant for each $k = 1, \ldots, n$.

Proof. By now the proof of this result should be clear to the reader. □

Theorem 7.2.3. *Let F be an algebraically closed field and let V be an n-dimensional vector space over F with $n \geq 1$. Then there is a basis $\mathcal{B}$ for V such that $[T]_\mathcal{B}$ is upper triangular.*

Proof. We use induction on the dimension of V. Clearly the theorem is true if $n = 1$. So suppose that $n > 1$ and that the theorem holds for all vector spaces over F whose diminsion is a positive integer less than n. Let λ be any eigenvalue of T (which we know must exist by Theorem 7.2.1). Let

$$U = \mathrm{Im}(T - \lambda I).$$

Because $T - \lambda I$ is not injective, it is also not surjective, so $\dim(U) < n = \dim(V)$. If $u \in U$, then

$$T(u) = (T - \lambda I)(u) + \lambda u.$$

Obviously $(T - \lambda I)(u) \in U$ (from the definition of U) and $\lambda u \in U$. Thus the equation above shows that $T(u) \in U$, hence U is T-invariant. Thus $T|_U \in \mathcal{L}(U)$. By our induction hypothesis, there is a basis $(u_1, \ldots, u_m)$ of U with respect to which $T|_U$ has an upper triangular matrix. Thus for each j we have (using Theorem 7.2.2)

$$T(u_j) = (T|_U)(u_j) \in span(u_1, \ldots, u_j). \tag{7.3}$$

Extend $(u_1, \ldots, u_m)$ to a basis $(u_1, \ldots, u_m, v_1, \ldots, v_r)$ of V. For each k, $1 \leq k \leq r$, we have

$$T(v_k) = (T - \lambda I)(v_k) + \lambda v_k.$$

The definition of U shows that $(T - \lambda I)(v_k) \in U = \text{span}(u_1, \ldots, u_m)$. Clearly

$$T(v_k) \in \text{span}(u_1, \ldots, u_m, v_1, \ldots, v_k).$$

It is now clear that that T has an upper triangular matrix with respect to the basis $(u_1, \ldots, u_m, v_1, \ldots, v_r)$. □

Obs. 7.2.4. *Suppose $T \in \mathcal{L}(V)$ has an upper triangular matrix with respect to some basis $\mathcal{B}$ of V. Then T is invertible if and only if all the entries on the diagonal of that upper triangular matrix are nonzero.*

Proof. We know that T is invertible if and only if $[T]_{\mathcal{B}}$ is invertible, which by Result 6.3.2 and Theorem 6.3.7 is invertible if and only if the diagonal entries of $[T]_{\mathcal{B}}$ are all nonzero. □

Corollary 7.2.5. *Suppose $T \in \mathcal{L}(V)$ has an upper triangular matrix with respect to some basis $\mathcal{B}$ of V. Then the eigenvalues of T consist precisely of the entries on the diagonal of $[T]_{\mathcal{B}}$.*

Proof. Suppose the diagonal entries of the $n \times n$ upper triangular matrix $[T]_{\mathcal{B}}$ are $\lambda_1, \ldots, \lambda_n$. Let $\lambda \in F$. Then $[T - \lambda I]_{\mathcal{B}}$ is upper triangular with diagonal elements equal to $\lambda_1 - \lambda, \lambda_2 - \lambda, \ldots, \lambda_n - \lambda$. Hence $T - \lambda I$ is not invertible if and only if λ equals one of the λ_j's. In other words, λ is an eigenvalue of T if and only λ equals one of the λ_j's as desired. □

Obs. 7.2.6. *Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a basis for V. An operator $T \in \mathcal{L}(V)$ has a diagonal matrix $diag(\lambda_1, \ldots, \lambda_n)$ with respect to $\mathcal{B}$ if and only if $T(v_i) = \lambda_i v_i$, i.e., each vector in $\mathcal{B}$ is an eigenvector of T.*

In some ways, the nicest operators are those which are diagonalizable, i.e., those for which there is some basis with respect to which they are represented by a diagonal matrix. But this is not always the case even when the field F is algebraically closed. Consider the following example over the complex numbers. Define $T \in \mathcal{L}(\mathcal{C}^2)$ by $T(y, z) = (z, 0)$. As you should verify, if $\mathcal{S}$ is the standard basis of $\mathcal{C}^2$, then $[T]_{\mathcal{S}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, so the only eigenvalue of T is 0. But $\text{null}(T - 0I)$ is 1-dimensional. So clearly $\mathcal{C}^2$ does not have a basis consisting of eigenvectors of T.

One of the recurring themes in linear algebra is that of obtaining conditions that guarantee that some operator have a diagonal matrix with respect to some basis.

Theorem 7.2.7. *If $dim(V) = n$ and $T \in \mathcal{L}(V)$ has exactly n distinct eigenvalues, then T has a diagonal matrix with respect to some basis.*

Proof. Suppose that T has distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, and let v_j be a nonzero eigenvector belonging to λ_j, $1 \leq j \leq n$. Because nonzero eigenvectors corresponding to distinct eigenvalues are linearly independent, $(v_1, \ldots, v_n)$ is linearly independent, and hence a basis of V. So with respect to this basis T has a diagonal matrix. □

The following proposition gathers some of the necessary and sufficient conditions for an operator T to be diagonalizable.

Theorem 7.2.8. *Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(i) T has a diagonal matrix with respect to some basis of V.

(ii) V has a basis consisting of eigenvectors of T.

(iii) There exist one-dimensional T-invariant subspaces $U_1, \ldots, U_n$ of V such that

$$V = U_1 \oplus \cdots \oplus U_n.$$

(iv) $V = null(T - \lambda_1 I) \oplus \cdots \oplus null(T - \lambda_m I)$.

(v) $dim(V) = dim(null(T - \lambda_1 I)) + \cdots dim(null(T - \lambda_m I))$.

Proof. At this stage it should be clear to the reader that (i), (ii) and (iii) are equivalent and that (iv) and (v) are equivalent. At least you should think about this until the equivalences are quite obvious. We now show that (ii) and (iv) are equivalent.

Suppose that V has a basis $\mathcal{B} = (v_1, \ldots, v_n)$ consisting of eigenvectors of T. We may group together those v_i's belonging to the same eigenvalue, say $v_1, \ldots, v_{d_1}$ belong to λ_1 so span a subspace U_1 of $\text{null}(T - \lambda_1 I)$; $v_{d_1+1}, \ldots, v_{d_1+d_2}$ belong to λ_2 so span a subspace U_2 of $\text{null}(T - \lambda_2 I)$; ..., and the last d_m of the v_i's belong to λ_m and span a subspace U_m of $\text{null}(T - \lambda_m I)$. Since $\mathcal{B}$ spans all of V, it is clear that $V = U_1 + \cdots + U_m$. Since the subspaces of eigenvectors belonging to distinct eigenvalues are independent (an easy corollary of Theorem 7.1.2), it must be that $V = U_1 \oplus \cdots \oplus U_m$. Hence each U_i is all of $\text{null}(T - \lambda_i I)$.

Conversely, if $V = U_1 \oplus \cdots \oplus U_m$, by joining together bases of the U_i's we get a basis of V consisting of eigenvectors of T. □

7.3 Invariant Subspaces of Real Vector Spaces

We have seen that if V is a finite dimensional vector space over an algebraically closed field F then each linear operator on V has an eigenvalue in F. We have also seen an example that shows that this is not the case for real vector spaces. This means that an operator on a nonzero finite dimensional real vector space may have no invariant subspace of dimension 1. However, we now show that an invariant subspace of dimension 1 or 2 always exists.

Theorem 7.3.1. *Every operator on a finite dimensional, nonzero, real vector space has an invariant subspace of dimension 1 or 2.*

Proof. Suppose V is a real vector space with $\dim(V) = n > 0$ and let $T \in \mathcal{L}(V)$. Choose $v \in V$ with $v \neq \vec{0}$. Since $(v, T(v), \ldots, T^n(v))$ must be linearly dependent (Why?), there are scalars $a_0, a_1, \ldots, a_n \in F$ such that not all the a_i's are zero and

$$\vec{0} = a_0 v + a_1 T(v) + \cdots a_n T^n(v).$$

Construct the polynomial $f(x) = \sum_{i=0}^{n} a_i x^i$ which can be factored in the form

$$f(x) = c(x - \lambda_1) \cdots (x - \lambda_r)(x^2 + \alpha_1 x + \beta_1) \cdots (x^2 + \alpha_k x + \beta_k),$$

where c is a nonzero real number, each λ_j, α_j, and β_j is real, $r + k \geq 1$, $\alpha_j^2 < 4\beta_j$ and the equation holds for all $x \in \mathcal{R}$. We then have

$$\begin{aligned} 0 &= a_0 v + a_1 T(v) + \cdots + a_n T^n(v) \\ &= (a_0 I + a_1 T + \cdots + a_n T^n)(v) \qquad (7.4) \\ &= c(T - \lambda_1 I) \cdots (T - \lambda_r I)(T^2 + \alpha_1 T + \beta_1 I) \cdots (T^2 + \alpha_k T + \beta_k I)(v), \end{aligned}$$

which means that $T - \lambda_j$ is not injective for at least one j or that $T^2 + \alpha_j T + \beta_j I$ is not injective for at least one j. If $T - \lambda_j I$ is not injective for some j, then T has an eigenvalue and hence a one-dimensional invariant subspace. If $T^2 + \alpha_j T + \beta_j I$ is not injective for some j, then we can find a nonzero vector v for which

$$T^2(v) + \alpha_j T(v) + \beta_j v = \vec{0}. \qquad (7.5)$$

Using Eq. 7.5 it is easy to show that $\text{span}(v, T(v))$ is T-invariant and it clearly has dimension 1 or 2. If it had dimension 1, then v would be an

eigenvector of T belonging to an eigenvalue λ that would have to be a root of $x^2 + \alpha_j x + \beta_j = 0$, contradicting the assumption that $\alpha_j^2 < 4\beta_j$. Hence T has a 2-dimensional invariant subspace. □

In fact we now have an easy proof of the following:

Theorem 7.3.2. *Every operator on an odd-dimensional real vector space has an eigenvalue.*

Proof. Let V be a real vector space with $\dim(V)$ odd and let $T \in \mathcal{L}(V)$. We know that the eigenvalues of T are the roots of the characteristic polynomial of T which has real coefficients and degree equal to $\dim(V)$. But every real polynomial of odd degree has a real root by Lemma 1.1.2, i.e., T has a real eigenvalue. □

7.4 Two Commuting Linear Operators*

Let V be a finite-dimensional vector space over the field F, and let T be a linear operator on V. Suppose T has matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ with respect to some basis. If i is an element in F (or in some extension of F) for which $i^2 = -1$, then the eigenvalues of T are $\pm i$. So T has eigenvectors in V if and only if $i \in F$. For example, if $F = \mathcal{R}$, then T has no eigenvectors. To avoid having to deal with this kind of situation we assume from now on that F *is algebraically closed*, so that each polynomial that has coefficients in F splits into linear factors over F. In particular, any linear operator T on V will have minimal and characteristic polynomials that split into linear factors over F. Our primary example of an algebraically closed field is the field $\mathcal{C}$ of complex numbers.

Recall that the ring $F[x]$ of polynomials in the indeterminate x with coefficients from F is a *principle ideal domain.* This means that if I is an *ideal* of $F[x]$, it must consist of all multiples of some particular element of $F[x]$. Our chief example is the following: Let W be a T-invariant subspace of V, T any linear operator on V, and let v be any vector in V. Put $T(v, W) = \{f(x) \in F[x] : f(T)(v) \in W\}$. It is easy to show that $T(v, W)$ is an ideal of $F[x]$. (This just means that the sum of any two polynomials in $T(v, W)$ is also in $T(v, W)$, and if $f(x) \in T(v, W)$ and $g(x)$ is any polynomial in $F[x]$, then the product $f(x)g(x)$ is back in $T(v, W)$.) Hence there is a unique

monic polynomial $g(x)$ of minimal degree in $T(v, W)$ called the *T-conductor of v into W*. For this conductor $g(x)$ it is true that $f(x) \in T(v, W)$ if and only if there is some $h(x) \in F[x]$ for which $f(x) = g(x) \cdot h(x)$. If $W = \{\overline{0}\}$, then $g(x)$ is called the *T-annihilator* of $\vec{v}$. Clearly the minimal polynomial $p(x)$ of T is in $T(v, W)$, so $g(x)$ *divides* $p(x)$. All these polynomials have coefficients in F, so by hypothesis they all split into linear factors over F.

The fact that $p(x)$ divides any polynomial $q(x)$ for which $q(T) = 0$ is quite important. Here is an example. Suppose W is a subspace invariant under T, so the restriction $T|_W$ of T to vectors in W is a linear operator on W. Let $g(x)$ be the minimal polynomial for $T|_W$. Since $p(T) = 0$, clearly $p(T|_W) = 0$, so $g(x)$ divides $p(x)$.

Theorem 7.4.1. *Let V be a finite-dimensional vector space over (the algebraically closed) field F, with $n = dim(V) \geq 1$. Let S and T be commuting linear operators on V. Then every eigenspace of T is invariant under S, and S and T have a common eigenvector in V.*

Proof. Since $dim(V) \geq 1$ and F is algebraically closed, T has an eigenvector v_1 in V with associated eigenvalue $c \in F$, i.e., $\overline{0} \neq v_1 \in V$ and $(T - cI)(v_1) = \overline{0}$. Put $W = \{v \in V : (T - cI)(v) = \overline{0}\}$, so W is the eigenspace associated with the eigenvalue c.

To see that W is invariant under S let $w \in W$, so $T(w) = cw$. Then since S commutes with T, it also commutes with $T - cI$, and $(T - cI)(S(w)) = [(T - cI)S](w) = [S(T - cI)](w) = S((T - cI)(w)) = S(\overline{0}) = \overline{0}$. This says that $S(w)$ is in W, so S acts on W, which has dimension at least 1. But then $S|_W$ is a linear operator on W and must have an eigenvector w_1 in W. So w_1 is a common eigenvector of S and T. □

Note: In the above proof we do not claim that every element of W is an eigenvector of S.

We need to use the concept of *quotient spaces*. Let W be a subspace of V. From Algebra we know that the quotient group V/W (considering the additive groups of V and W) is also an abelian group. We can make it into a vector space over the same field by defining a scalar multiplication as follows: for $c \in F$ and $v + W \in V/W$, put $c(v + W) = cv + W$. It is routine to show that this makes V/W into a vector space. Moreover, if $\mathcal{B}_1 = (v_1, \ldots, v_r)$ is a basis for W, and $\mathcal{B}_2 = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ is a basis for V, then $(v_{r+1} + W, \ldots, v_n + W)$ is a basis for V/W. Hence $\dim(V) = \dim(W) +$

$\dim(V/W)$. Moreover, if W is invariant under the operator $T \in \mathcal{L}(V)$, then T induces a linear operator $\overline{T}$ on V/W as follows:

$$\overline{T} : V/W \to V/W : v + W \mapsto T(v) + W.$$

Clearly if $T, S \in \mathcal{L}(V)$ and $S \cdot T = T \cdot S$, then $\overline{T} \cdot \overline{S} = \overline{S} \cdot \overline{T}$.

We are now ready for the following theorem on two commuting operators.

Theorem 7.4.2. *Let S and T be two commuting operators on V over the algebraically closed field F. Then there is a basis $\mathcal{B}$ of B with respect to which both $[T]_{\mathcal{B}}$ and $[S]_{\mathcal{B}}$ are upper triangular.*

Proof. We proceed by induction on n, the dimension of V.

By the previous theorem there is a vector $v_1 \in V$ such that $Tv_1 = \lambda_1 v_1$ and $Sv_1 = \mu_1 v_1$ for some scalars λ_1 and μ_1. Let W be the subspace spanned by v_1. Then the dimension of V/W is $n - 1$, and the operators $\overline{T}$ and $\overline{S}$ on V/W commute, so by our induction hypothesis there is a basis $\mathcal{B}_1 = (v_2 + W, v_3 + W, \ldots, v_n + W)$ of V/W with respect to which both $\overline{T}$ and $\overline{S}$ have upper triangular matrices. It follows that $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ is a basis of V with respect to which both T and S have upper triangular matrices. □

Theorem 7.4.3. *Let T be a diagonalizable operator on V and let $S \in \mathcal{L}(V)$. Then $ST = TS$ if and only if each eigenspace of T is invariant under S.*

Proof. Since T is diagonalizable, there is a basis $\mathcal{B}$ of V consisting of eigenvectors of T. Let W be the eigenspace of T associated with the eigenvalue λ, and let $w \in W$. If $S(w) \in W$, then $(TS)(w) = T(Sw) = \lambda Sw = S(\lambda w) = S(Tw) = ST(w)$. So if each eigenspace of T is invariant under S, S and T commute at each element of $\mathcal{B}$, implying that $ST = TS$ on all of V. Conversely, suppose that $ST = TS$. Then for any $w \in W$, $T(Sw) = S(Tw) = S(\lambda w) = \lambda S(w)$, implying that $Sw \in W$. So each eigenspace of T must be invariant under S. □

Note that even if T is diagonalizable and S commutes with T, it need not be the case that S must be diagonalizable. For example, if $T = I$, then V is the only eigenspace of T, and if S is any non-diagonalizable operator on V, then S still commutes with T. However, if both T and S are known to be diagonalizable, then we can say a bit more.

Theorem 7.4.4. *Let S and T both be diagonalizable operators on the n-dimensional vector space V over the field F. Then S and T commute if and only if S and T are simultaneously diagonalizable.*

Proof. First suppose that S and T are simultaneously diagonalizable. Let $\mathcal{B}$ be a basis of V with respect to which both $[T]_{\mathcal{B}}$ and $[S]_{\mathcal{B}}$ are diagonal. Since diagonal matrices commute, $[T]_{\mathcal{B}}$ and $[S]_{\mathcal{B}}$ commute, implying that T and S commute.

Conversely, suppose that T and S commute. We proceed by induction on n. If $n = 1$, then any basis is equivalent to the basis consisting of any nonzero vector, and any 1×1 matrix is diagonal. So assume that $1 < n$ and the result holds over all vector spaces of dimension less than n. If T is a scalar times the identity operator, then clearly any basis that diagonalizes S will diagonalize both S and T. So suppose T is not a scalar times the identity. Let λ be an eigenvalue of T and put $W = \text{null}(T - \lambda I)$. So W is the eigenspace of T associated with λ, and by hypothesis $1 \leq \dim(W) < n$. By Theorem 7.4.3 W is invariant under S. It follows that $T|_W$ and $S|_W$ are simultaneously diagonalizable. Let $\mathcal{B}_W$ be a basis for W which consists of eigenvectors of both $T|_W$ and $S|_W$, so they are also eigenvectors of both T and S. Repeat this process for each eigenvalue of T and let $\mathcal{B}$ be the union of all the bases of the various eigenspaces. Then the matrices $[T]_{\mathcal{B}}$ and $[S]_{\mathcal{B}}$ are both diagonal. □

Theorem 7.4.5. *Let T be a diagonalizable operator on V, an n-dimensional vector space over F. Let $S \in \mathcal{L}(V)$. Then there is a polynomial $f(x) \in F[x]$ such that $S = f(T)$ if and only if each eigenspace of T is contained in a single eigenspace of S.*

Proof. First suppose that $S = f(T)$ for some $f(x) \in F[x]$. If $T(v) = \lambda v$, then $S(v) = f(T)(v) = f(\lambda)v$. Hence the entire eigenspace of T associated with λ is contained in the eigenspace of S associated with its eigenvalue $f(\lambda)$. This completes the proof in one direction.

For the converse, let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues of T, and let $W_i = \text{null}(T - \lambda_i I)$, the eigenspace of T associated with λ_i. Let μ_i be the eigenvalue of S whose corresponding eigenspace contains W_i. Note that the values $\mu_1, \ldots, \mu_r$ might not be distinct. Use Lagrange interpolation to construct the polynomial $f(x) \in F[x]$ for which $f(\lambda_i) = \mu_i$, $1 \leq i \leq r$. Then define an operator $S' \in \mathcal{L}(V)$ as follows. Let $\mathcal{B}$ be a basis of V consisting of

the union of bases of the eigenspaces W_i. For each $v \in \mathcal{B}$, say $v \in W_i$, put $S'(v) = f(T)(v) = f(\lambda_i)v = \mu_i v = S(v)$. Then since S' and S agree on a basis of V, they must be the same operator. Hence $S = f(T)$. □

Corollary 7.4.6. *Let $T \in \mathcal{L}(V)$ have n distinct eigenvalues where $n = dim(V)$. Then the following are equivalent:*
(i) $ST = TS$.
(ii) S and T are simultaneously diagonalizable.
(ii) S is a polynomial in T.

Proof. Since T has n distinct eigenvalues, its minimal polynomial has no repeated factors, so T is diagonalizable. Then Theorem 7.4.3 says that $ST = TS$ iff each eigenspace of T is invariant under S. Since each eigenspace of T is 1-dimensional, this means that each eigenvector of T is also an eigenvector of S. Using Theorem 7.4.5 we easily see that the theorem is completely proved. □

We note the following example: If T is any invertible operator, so the constant term of its minimal polynomial is not zero, it is easy to use the minimal polynomial to write I as a polynomial in T. However, any polynomial in I is just some constant times I. So if T is any invertible operator that is not a scalar times I, then I is a polynomial in T, but T is not a polynomial in I.

7.5 Commuting Families of Operators*

Let V be an n-dimensional vector space over F, and let $\mathcal{F}$ be a family of linear operators on V. We want to know when we can simultaneously triangulate or diagonalize the operators in $\mathcal{F}$, i.e., find one basis $\mathcal{B}$ such that all of the matrices $[T]_{\mathcal{B}}$ for $T \in \mathcal{F}$ are upper triangular or diagonal. In the case of diagonalization, it is necessary that F be a commuting family of operators: $UT = TU$ for all $T, U \in \mathcal{F}$. That follows from the fact that all diagonal matrices commute. Of course, it is also necessary that each operator in $\mathcal{F}$ be a diagonalizable operator. In order to simultaneously triangulate, each operator in $\mathcal{F}$ must be triangulable. It is not necessary that $\mathcal{F}$ be a commuting family; howefver, that condition is sufficient for simultaneous triangulation as long as each T in $\mathcal{F}$ can be individually triangulated.

The subspace W of V is *invariant under* (the family of operators) $\mathcal{F}$ if W is invariant under each operator in $\mathcal{F}$.

Lemma 7.5.1. *Let $\mathcal{F}$ be a commuting family of triangulable linear operators on V. Let W be a proper subspace of V which is invariant under $\mathcal{F}$. There exists a vector $v \in V$ such that:*

(a) v is not in W;

(b) for each T in $\mathcal{F}$, the vector $T(v)$ is in the subspace spanned by v and W.

Proof. Since the space of all linear operators on V is a vector space with dimension n^2, it is easy to see that without loss of generality we may assume that $\mathcal{F}$ has only finitely many operators. For let $\{T_1, \ldots, T_r\}$ be a maximal linearly independent subset of $\mathcal{F}$, i.e., a basis for the subspace of $\mathcal{L}(V)$ spanned by $\mathcal{F}$. If v is a vector such that (b) holds for each T_i, then (b) holds for each each operator which is a linear combination of $T_1, \ldots, T_r$.

First we establish the lemma for one operator T. To do this we need to show that the T-conductor of v into W is a linear polynomial. Since T is triangulable, its minimal polynomial $p(x)$, as well as its characteristic polynomial, factor over F into a product of linear factors. Say $p(x) = (x - c_1)^{e_1}(x - c_2)^{e_2} \cdots (x - c_r)^{e_r}$. Let w be any vector of V that is not in W, and let $g(x)$ be the T-conductor of w into W. Then g divides the minimal polynomial for T. Since w is not in W, g is not constant. So $g(x) = (x - c_1)^{f_1}(x - c_2)^{f_2} \cdots (x - c_r)^{f_r}$ where at least one of the integers f_i is positive. Choose j so that $f_j > 0$. Then $(x - c_j)$ divides $g(x)$:

$$g = (x - c_j)h.$$

By definition of g, the vector $u = h(T)(w)$ cannot be in W. But

$$\begin{aligned} (T - c_jI)(u) &= (T - c_jI)h(T)(w) \\ &= g(T)(w) \in W. \end{aligned} \tag{7.6}$$

Now return to thinking about the family $\mathcal{F}$. By the previous paragraph we can find a vector v_1 not in W and a scalar c_1 such that $(T_1 - c_1I)(v_1) \in W$. Let V_1 be the set of all vectors $v \in V$ such that $(T_1 - c_1I)(v) \in W$. Then V_1 is a subspace of V that properly contains W. Since $T \in \mathcal{F}$ commutes with T_1 we have

$$(T_1 - c_1I)(T(v)) = T(T_1 - c_1I)(v).$$

If $v \in V_1$, then $(T_1 - c_1 I)v \in W$. Since W is invariant under each T in $\mathcal{F}$, we have $T(T_1 - c_1 I)(v) \in W$ i.e., $Tv \in V_1$, for all $v \in V_1$ and all $T \in \mathcal{F}$.

Note again that W is a proper subspace of V_1. Put $U_2 = T_2|_{V_1}$, the operator obtained by restricting T_2 to the subspace V_1. The minimal polynomial for U_2 divides the munimum polynomial for T_2. By the second paragraph of this proof applied to U_2 and the invariant subspace W, there is a vector $v_2 \in V_1$ but not in W, and a scalar c_2 such that $(T_2 - c_2 I)(v_2) \in W$. Note that

(a) $v_2 \notin W$;
(b) $(T_1 - c_1 I)(v_2) \in W$;
(c) $(T_2 - c_2 I)(v_2) \in W$.

Let $V_2 = \{v \in V_1 : (T_2 - c_2 I)(v) \in W\}$. Then V_2 is invariant under $\mathcal{F}$. Apply the same ideas to $U_3 = T_3|_{V_2}$. Continuing in this way we eventually find a vector $v = v_r$ not in W such that $(T_j - c_j I)(v) \in W$, for $1 \leq j \leq r$. □

Theorem 7.5.2. *Let V be a finite-dimensional vector space over the field F. Let $\mathcal{F}$ be a commuting family of triangulable linear operators on V (i.e., the minimal polynomial of each $T \in \mathcal{F}$ splits into linear factors). There exists an ordered basis for V such that every operator in $\mathcal{F}$ is represented by a triangular matrix with respect to that basis.*

Proof. Start by applying Lemma 7.5.1 to the $\mathcal{F}$-invariant subspace $W = \{\overline{0}\}$ to obtain a nonzero vector v_1 for which $T(v_1) \in W_1 = \langle v_1 \rangle$ for all $T \in \mathcal{F}$. Then apply the lemma to W_1 to find a vector v_2 not in W_1 but for which $T(v_2) \in W_2 = \langle v_1, v_2 \rangle$. Proceed in this way until a basis $\mathcal{B} = (v_1, v_2, \ldots)$ of V has been obtained. Clearly $[T]_{\mathcal{B}}$ is upper triangular for every $T \in \mathcal{B}$. □

Corollary 7.5.3. *Let $\mathcal{F}$ be a commuting family of $n \times n$ matrices over an algebraically closed field F. There exists a nonsingular $n \times n$ matrix P with entries in F such that $P^{-1}AP$ is upper-triangular, for every matrix A in $\mathcal{F}$.*

Theorem 7.5.4. *Let $\mathcal{F}$ be a commuting family of diagonalizable linear operators on the finite-dimensional vector space V. There existws an ordered basis for V such that every operator in $\mathcal{F}$ is represented in that basis by a diagonal matrix.*

Proof. If $\dim(V) = 1$ or if each $T \in \mathcal{F}$ is a scalar times the identity, then there is nothing to prove. So suppose $1 < n = \dim(V)$ and that the theorem is true for vector spaces of diimension less than n. Also assume that for

some $T \in \mathcal{F}$, T is not a scalar multiple of the identity. Since the operators in $\mathcal{F}$ are all diagonalizable, we know that each minimal polynomial splits into distinct linear factors. Let $c_1, \ldots, c_k$ be the distinct eigenvalues of T, and for each index i put $W_i = \text{null}(T - c_i I)$. Fix i. Then W_i is invariant under every operator that commutes with T. Let $\mathcal{F}_i$ be the family of linear operators on W_i obtained by restricting the operators in $\mathcal{F}$ to the invariant subspace W_i. Each operator in $\mathcal{F}_i$ is diagonalizable, because its minimal polynomial divides the minimal polynomial for the corresponding operator in $\mathcal{F}$. By hypothesis, $\dim(W_i) < \dim(V)$. So the operators in $\mathcal{F}_i$ can be simultaneously diagonalized. In other words, W_i has a basis $\mathcal{B}_i$ which consists of vectors which are simultaneously characteristic vectors for every operator in $\mathcal{F}_i$. Then $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is the basis of V that we seek. □

7.6 The Fundamental Theorem of Algebra*

This section is based on the following article: Harm Derksen, The Fundamental Theorem of Algebra and Linear Algebra, *The American Mathematical Monthly*, vol 110, Number 7, August-September 2003, 620 – 623. We start by quoting from the third paragraph of the article by Derksen.

"Since the fundamental theorem of algebra is needed in linear algebra courses, it would be desirable to have a proof of it in terms of linear algebra. In this paper we prove that every square matrix with complex coefficients has an eigenvector. This statement is equivalent to the fundamental theorem of algebra. In fact, we will prove the slightly stronger result that any number of commuting square matrices with complex entries have a common eigenvector. The proof lies entirely within the framework of linear algebra, and unlike most other algebraic proofs of the fundamental theorem of algebra, it does not require Galois theory or splitting fields. "

Preliminaries

Several results we have obtained so far have made the assumption that the field F was algebraically closed. Moreover, we often gave the complex numbers $\mathcal{C}$ as the prototypical example. So in this section we have to be careful not to quote any results that might have hidden in them the assumption that $\mathcal{C}$ is algebraically closed.

For the proof we use only the following elementary properties of real and complex numbers that were established much earlier.

Lemma Every polynomial of odd degree with real coefficients has a (real) zero.

Lemma Every complex number has a square root.

Theorem 7.3.2 If A is real, $n \times n$, with n odd, then A has an eigenvector (belonging to a real eigenvalue).

An Induction Argument

Keep in mind that we cannot use results that might have hidden in them the assumption that $\mathcal{C}$ is algebraically closed.

For a field K and for positive integers d and r, consider the following statement:

$P(K, d, r)$: Any r commuting linear transformations $A_1, A_2, \ldots, A_r$ of a K- vector space V of dimension n such that d does not divide n have a common eigenvector.

Lemma 7.6.1. *If $P(K, d, 1)$ holds, then $P(K, d, r)$ holds for all $r \geq 1$.*

It is important to realize that the smallest d for which the hypothesis of this lemma holds in a given situation might be much larger than $d = 1$.

Proof. The proof is by induction on r. The case of $P(K, d, 1)$ is true by hypothesis. For $r \geq 2$, suppose that $P(K, d, r-1)$ is true and let $A_1, \ldots, A_r$ be commuting linear transformations of V of dimension such that d does not divide n. Because $(K, d, 1)$ holds, A_r has an eigenvalue λ in K. Let W be the kernel and Z the image of $A_r - \lambda I$. It is now easy to show that each of W and Z are left invariant by each of $A_1, \ldots, A_{r-1}$.

First suppose that $W \neq V$. Because $\dim W + \dim Z = \dim V$, either d does not divide $\dim W$ or d does not divide $\dim Z$. Since $\dim W < n$ and $\dim Z < n$, we may assume by induction on n that $A_1, \ldots, A_r$ already have a common eigenvector in W or in Z.

In the remaining case, $W = V$. Because $P(K, d, r-1)$ holds, we may assume that $A_1, \ldots, A_{r-1}$ have a common eigenvector in V, say v. Since $A_r v = \lambda v$ (because $W = V$), v is a common eigenvector of $A_1, \ldots, A_r$. $\square$

Lemma 7.6.2. *$P(K, 2, r)$ holds for all $r \geq 1$. In other words, if $A_1, \ldots, A_r$ are commuting linear transformations on an odd dimensional $\mathcal{R}$-vector space, then they have a common eigenvector.*

Proof. By Lemma 7.6.1 it is enough to show that $P(\mathcal{R}, 2, 1)$ is true. If A is an linear transformation of an odd dimensional $\mathcal{R}$-vector space, $\det(xI - A)$ is a polynomial of odd degree, which has a zero λ by Lemma 1.1.2. Then λ is a real eigenvalue of A. □

We now "lift" the result of Lemma 7.6.2 to the analogous result over the field $\mathcal{C}$.

Definition If A is any $m \times n$ matrix over $\mathcal{C}$ we let A^* denote the transpose of the complex conjugate of the matrix A, i.e., $(A^*)_{ij} = \overline{A}_{ij}$.

Lemma 7.6.3. *$P(\mathcal{C}, 2, 1)$ holds, i.e., every linear transformation of a $\mathcal{C}$-vector space of odd diimension has an eigenvector.*

Proof. Suppose that $A : \mathcal{C}^n \to \mathcal{C}^n$ is a $\mathcal{C}$-linear map with n odd. Let V be the $\mathcal{R}$-vector space $\mathrm{Herm}_n(\mathcal{C})$, the set of $n \times n$ Hermitian matrices. Define two linear operators L_1 and L_2 on V by

$$L_1(B) = \frac{AB + BA^*}{2},$$

and

$$L_2(B) = \frac{AB - BA^*}{2i}.$$

It is now easy to show that $\dim V = n^2$, which is odd. It is also routine to check that L_1 and L_2 commute. Hence by Lemma 7.6.2, $P(\mathcal{R}, 2, 2)$ holds and implies that L_1 and L_2 have a common eigenvector B, say $L_1(B) = \lambda B$ and $L_2(B) = \mu B$, with λ and μ both real. But then

$$(L_1 + iL_2)(B) = AB = (\lambda + \mu i)B,$$

and any nonzero column vector of B gives an eigenvector for the matrix A. □

Lemma 7.6.4. *$P(\mathcal{C}, 2^k, r)$ holds for all $k \geq 1$ and $r \geq 1$.*

Proof. The proof is by induction on k. The case $k = 1$ follows from Lemmas 7.6.3 and 7.6.1. Assume that $P(\mathcal{C}, 2^l, r)$ holds for $l < k$. We will establish that $P(\mathcal{C}, 2^k, r)$ holds. In view of Lemma 7.6.1 it suffices to prove $P(\mathcal{C}, 2^k, 1)$. Suppose that $A : \mathcal{C}^n \to \mathcal{C}^n$ is linear, where n is divisible by 2^{k-1} by not by 2^k. Let V be the $\mathcal{C}$-vector space $\mathrm{Skew}_n(\mathcal{C}) = \{B \in M_n(\mathcal{C}) : B^* = -B\}$, the set of $n \times n$ skew-symmetric matrices with complex entries. Define two commuting linear transformations L_1 and L_2 of V by

$$L_1(B) = AB - BA^T$$

and

$$L_2(B) = ABA^T.$$

Note that $\dim V = n(n-1)/2$, which ensures that 2^{k-1} does not divide $\dim V$. By $P(\mathcal{C}, 2^{k-1}, 2)$, L_1 and L_2 have a common eigenvector B, say $L_1(B) = \lambda(B)$ and $L_2(B) = \mu B$, where λ and μ are now complex numbers. It follows that

$$\mu B = ABA^T = A(AB - \lambda B),$$

so

$$(A^2 - \lambda A - \mu I)B = 0.$$

Let v be a nonzero column of B. Then

$$(A^2 - \lambda A - \mu I)v = 0.$$

Since each element of $\mathcal{C}$ has a square root, there is a δ in $\mathcal{C}$ such that $\delta^2 = \lambda^2 + 4\mu$. We can write $x^2 - \lambda x - \mu = (x-\alpha)(x-\beta)$, where $\alpha = (\lambda+\delta)/2$ and $\beta = (\lambda - \delta)/2$. We then have

$$(A - \alpha I)w = 0,$$

where $w = (A - \beta I)v$. If $w = 0$, then v is an eigenvector of A with eigenvalue β; if $w \neq 0$, then w is an eigenvector of A with eigenvalue α. □

We have now reached the point where we can prove the main result for commuting operators on a complex space.

Theorem 7.6.5. *If $A_1, A_2, \ldots, A_r$ are commuting linear transformations of a finite dimensional nonzero $\mathcal{C}$-vector space V, then they have a common eigenvector.*

Proof. Let n be the dimension of V. There exists a positive integer k such that 2^k does not divide n. Since $P(\mathcal{C}, 2^k, r)$ holds by Lemma 7.6.4, the theorem follows. □

Theorem 7.6.6. *(The Fundamental Theorem of Algebra) If $P(x)$ is a non-constant polynomial with complex coefficients, then there exists a λ in $\mathcal{C}$ such that $P(\lambda) = 0$.*

Proof. It suffices to prove this for monic polynomials. So let

$$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n.$$

Then $P(x) = \det(xI - A)$, where A is the companion matrix of P:

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}.$$

Theorem 7.6.5 implies that A has a complex eigenvalue λ in $\mathcal{C}$, from which it follows that $P(\lambda) = 0$. □

7.7 Exercises

1. Let V be finite dimensional over F and let $P \in \mathcal{L}(V)$ be idempotent. Determine the eigenvalues of P and show that P is diagonalizable.

2. If $T, S \in \mathcal{L}(V)$ and $TS = ST$, show that

   (i) null(T) is S-invariant; and

   (ii) If $f(x) \in F[x]$, then null$(f(T))$ is S-invariant.

3. Suppose n is a positive integer and $T \in \mathcal{L}(F^n)$ is defined by

   $$T(z_1, z_2, \ldots, z_n) = (z_1 + \cdots + z_n, z_1 + \cdots + z_n, \ldots, z_1 + \cdots + z_n).$$

   Determine all eigenvalues and eigenvectors of T.

4. Suppose $T \in \mathcal{L}(V)$ and $\dim(\mathrm{Im}(T)) = k$. Prove that T has at most $k + 1$ distinct eigenvalues.

5. Suppose that $S, T \in \mathcal{L}(V)$, $\lambda \in F$ and $1 \leq k \in \mathcal{Z}$. Show that $(TS - \lambda I)^k T = T(ST - \lambda I)^k$.

6. Suppose that $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues but not necessarily the same minimal polynomial.

7. Suppose that $S, T \in \mathcal{L}(V)$ and at least one of S, T is invertible. Show that ST and TS have the same minimal and characteristic polynomials.

8. Suppose that F is algebraically closed, $p(z) \in F[z]$ and $a \in F$. Prove that a is an eigenvalue of $p(T)$ if and only if $a = p(\lambda)$ for some eigenvalue λ of T. (Hint: Suppose that a is an eigenvalue of $p(T)$. Factor $p(z) - a = c(z - \lambda_1) \cdots (z - \lambda_m)$. Use the fact that $p(T) - aI$ is not injective. Don't forget to consider what happens if $c = 0$.)

9. Suppose that $S, T \in \mathcal{L}(V)$ and that T is diagonalizable. Suppose that each eigenvector of T is an eigenvector of S. Show that $ST = TS$.

Chapter 8

Inner Product Spaces

8.1 Inner Products

Throughout this chapter F will denote a subfield of the complex numbers $\mathcal{C}$, and V will denote a vector space over F.

Definition An *inner product* on V is a scalar-valued function $\langle\ ,\ \rangle : V \times V \to F$ that satisfies the following properties:

(i) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(ii) $\langle cu, v\rangle = c\langle u, v\rangle$ for all $c \in F$, $u, v \in V$.

(iii) $\langle v, u\rangle = \overline{\langle u, v\rangle}$ for all $u, v \in V$, where the overline denotes complex conjugate .

(iv) $\langle u, u\rangle > 0$ if $u \neq \vec{0}$.

It is easy to check that the above properties force

$$\langle u, cv + w\rangle = \overline{c}\langle u, v\rangle + \langle u, w\rangle \quad \forall u, v, w \in V,\ c \in F.$$

Example 8.1.1. *On F^n there is a "standard" inner product defined as follows: for $\vec{x} = (x_1, \ldots, x_n)$ and $\vec{y} = (y_1, \ldots, y_n)$, put*

$$\langle \vec{x}, \vec{y}\rangle = \sum_{i=0}^{n} x_i \overline{y_i}.$$

It is easy enough to show that the above definition really does give an inner product on F^n. In fact, it is a special case of the following example.

Example 8.1.2. *Let A be an invertible $n \times n$ matrix over F. Define $\langle\ ,\ \rangle$ on F^n (whose elements are written as row vectors for the purpose of this example) as follows. For $u, v \in V$ put*

$$\langle u, v\rangle = uAA^*v^*, \quad \text{where } B^* \text{ denotes the complex conjugate transpose of } B.$$

It is a fairly straightforward exercise to show that this definition really gives an inner product. If $A = I$ the standard inner product of the previous example is obtained.

Example 8.1.3. *Let $C(0,1)$ denote the vector space of all continuous, real-valued functions on the interval $[0,1]$. For $f, g \in C(0,1)$ define*

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt.$$

Again it is a routine exercise to show that this really gives an inner product.

Definition An *inner product space* is a vector space over F (a subfield of $\mathcal{C}$) together with a specified inner product on that space.

Let V be an inner product space with inner product $\langle\ ,\ \rangle$. The length of a vector $v \in V$ is defined to be $||v|| = \sqrt{\langle v, v\rangle}$.

Theorem 8.1.4. *If V is an inner product space, then for any vectors $u, v \in V$ and any scalar $c \in F$,*
(i) $||cu|| = |c|\ ||u||$;
(ii) $||u|| > 0$ for $u \neq \vec{0}$;
(iii) $|\langle u, v\rangle| \leq ||u||\ ||v||$;
(iv) $||u + v|| \leq ||u|| + ||v||$.

Proof. Statements (i) and (ii) follow almost immediately from the various definitions involved. The inequality in (iii) is clearly valid when $u = \vec{0}$. If $u \neq \vec{0}$, put

$$w = v - \frac{\langle v, u\rangle}{||u||^2}u.$$

It is easily checked that $\langle w, u\rangle = 0$ and

$$\begin{aligned} 0 \leq ||w||^2 &= \langle v - \frac{\langle v,u\rangle}{||u||^2}u, v - \frac{\langle v,u\rangle}{||u||^2}u\rangle \\ &= \langle v,v\rangle - \frac{\langle v,u\rangle\langle u,v\rangle}{||u||^2} \\ &= ||v||^2 - \frac{|\langle u,v\rangle|^2}{||u||^2}. \end{aligned}$$

Hence $|\langle u,v\rangle|^2 \leq ||u||^2||v||^2$. It now follows that

$$\mathrm{Re}\langle u,v\rangle \leq |\langle u,v\rangle| \leq \|u\| \cdot \|v\|,$$

and

$$\begin{aligned} ||u+v||^2 &= ||u||^2 + \langle u,v\rangle + \langle v,u\rangle + ||v||^2 \\ &= ||u||^2 + 2\mathrm{Re}\langle u,v\rangle + ||v||^2 \\ &\leq ||u||^2 + 2||u||\ ||v|| + ||v||^2 \\ &= (||u|| + ||v||)^2. \end{aligned}$$

Thus $||u+v|| \leq ||u|| + ||v||$. □

The inequality in (iii) is called the **Cauchy-Schwarz inequality**. It has a very wide variety of applications. The proof shows that if u is nonzero, then $|\langle u,v\rangle| < ||u||\ ||v||$ unless

$$v = \frac{\langle v,u\rangle}{||u||^2}u,$$

which occurs if and only if (u,v) is a linearly dependent list. You should try out the Cauchy-Schwarz inequality on the examples of inner products given above. The inequality in (iv) is called the **triangle inequality**.

Definitions Let u and v be vectors in an inner product space V. Then u is *orthogonal to* v if and only if $\langle u,v\rangle = 0$ if and only if $\langle v,u\rangle = 0$, in which case we say u and v are orthogonal and write $u \perp v$. If S is a set of vectors in V, S is called an *orthogonal set* provided each pair of distinct vectors in S is orthogonal. An *orthonormal set* is an orthogonal set S with the additional property that $||u|| = 1$ for every $u \in S$. Analogous definitions are made for lists of vectors.

Note: The standard basis of F^n is an orthonormal list with respect to the standard inner product.

Also, the zero vector is the only vector orthogonal to every vector. (Prove this!)

Theorem 8.1.5. *An orthogonal set of nonzero vectors is linearly independent.*

Proof. Let S be a finite or infinite orthogonal set of nonzero vectors in a given inner product space. Suppose $v_1, \ldots, v_m$ are distinct vectors in S and that

$$w = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m.$$

Then

$$\begin{aligned} \langle w, v_k \rangle &= \langle \sum_j c_j v_j, v_k \rangle \\ &= \sum_j c_j \langle v_j, v_k \rangle \\ &= c_k \langle v_k, v_k \rangle. \end{aligned}$$

Since $\langle v_k, v_k \rangle \neq 0$, it follows that

$$c_k = \frac{\langle w, v_k \rangle}{||v_k||^2}, \quad 1 \leq k \leq m.$$

When $w = \vec{0}$, each $c_k = 0$, so S is an independent set. □

Corollary 8.1.6. *If a vector w is a linear combination of an orthogonal list $(v_1, \ldots, v_m)$ of nonzero vectors, then w is the particular linear combination*

$$w = \sum_{k=1}^{m} \frac{\langle w, v_k \rangle}{||v_k||^2} v_k. \tag{8.1}$$

Theorem 8.1.7. *(***Pythagorean Theorem***) If $u \perp v$, then*

$$||u + v||^2 = ||u||^2 + ||v||^2. \tag{8.2}$$

Proof. Suppose $u \perp v$. Then

$$\begin{aligned} ||u + v||^2 &= \langle u + v, u + v \rangle \\ &= ||u||^2 + ||v||^2 + \langle u, v \rangle + \langle v, u \rangle \\ &= ||u^2|| + ||v||^2. \end{aligned}$$

□

Theorem 8.1.8. *(***Parallelogram Equality***) If u, v are vectors in the inner product space V, then*

$$||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2).$$

Proof. The details are routine and are left as an exercise (see exercise 1). □

Starting with an inner product $\langle\ ,\ \rangle$ on a vector space U over F (where F is some subfield of $\mathcal{C}$), we defined a *norm* on U by $||v|| = \sqrt{\langle v, v\rangle}$, for all $v \in U$. This norm function satisfies a variety of properties as we have seen in this section. Sometimes we have a norm function given and would like to know if it came from an inner product. The next theorem provides an answer to this question, but first we give an official definition of a norm.

Definition A *norm* on a vector space U over the field F is a function

$$||\ \ || : U \to [0, \infty) \subseteq \mathcal{R} \text{ such that}$$

(i) $||u|| = 0$ iff $u = \vec{0}$;
(ii))$||au|| = |a| \cdot ||u||\ \forall a \in F,\ u \in U$;
(iii) $||u+v|| \leq ||u|| + ||v||$.

Theorem 8.1.9. *Let $||\ \ ||$ be a norm on U. Then there is an inner product $\langle\ ,\ \rangle$ on U such that $||u|| = \langle u, u\rangle^{\frac{1}{2}}$ for all $u \in U$ if and only if $||\ \ ||$ satisfies the parallelogram equality.*

Proof. We have already seen that if the norm is derived from an inner product, then it satisfies the parallelogram equality. For the converse, now suppose that $||\ \ ||$ is a norm on U satisfying the parallelogram equality. We will show that there must have been an inner product from which the norm was derived in the usual fashion. We first consider the case $F = \mathcal{R}$. It is then clear (from the real polarization identity - see Exercise 10) that $\langle\ ,\ \rangle$ must be defined in the following way:

$$\langle u, v\rangle = \frac{||u+v||^2 - ||u-v||^2}{4}. \tag{8.3}$$

It is then clear that $\langle u, u\rangle = \frac{||2u||^2 - |\vec{0}||^2|}{4} = ||u||^2$, so $||u|| = \langle u, u\rangle^{\frac{1}{2}}$ for all $u \in U$. but it is not at all clear that $\langle\ ,\ \rangle$ is an inner product. However, since $\langle u, u\rangle = ||u||^2$, by the definition of norm we see that

(a) $\langle u, u\rangle \geq 0$, with equality if and only if $u = \vec{0}$.

Next we show that $\langle\ ,\ \rangle$ is additive in the first slot. We use the parallelogram equality in the form $||u||^2 + ||v||^2 = \frac{||u+v||^2}{2} + \frac{||u-v||^2}{2}$.

Let $u, v, w \in U$. Then from the definition of $\langle\ ,\ \rangle$ we have:

$$4(\langle u+v, w\rangle - \langle u, w\rangle - \langle v, w\rangle \quad \text{(which should be 0)}$$

$$= ||u+v+w||^2 - ||u+v-w||^2 - ||u+w||^2 + ||u-w||^2 - ||v+w||^2 + ||v-w||^2$$

$$= ||u+v+w||^2 + (||u-w||^2 + ||v-w||^2) - ||u+v-w||^2 - (||u+w||^2 + ||v+w||^2)$$

$$= ||u+v+w||^2 + \frac{||u+v-2w||^2}{2} + \frac{||u-v||^2}{2} - ||u+v-w||^2 - \frac{u+v+2w||^2}{2} - \frac{||u-v||^2}{2}$$

$$= (||u+v+w||^2 + ||w||^2) + \frac{||u+v-2w||^2}{2} - (||u+v-w||^2 + ||w||^2) - \frac{||u+v+2w||^2}{2}$$

$$= \frac{||u+v+2w||^2}{2} + \frac{||u+v||^2}{2} + \frac{||u+v-2w||^2}{2}$$

$$- \frac{||u+v||^2}{2} - \frac{||u+v-2w||^2}{2} - \frac{||u+v+2w||^2}{2}$$

$$= 0.$$

Hence $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$, proving

(b) $\langle\ ,\ \rangle$ is additive in the first slot.

To prove that $\langle\ ,\ \rangle$ is homogeneous in the first slot is rather more involved. First suppose that n is a positive integer. Then using additivity in the first slot we have $\langle nu, v\rangle = n\langle u, v\rangle$. Replacing u with $\frac{1}{n}u$ gives $\langle u, v\rangle = n\langle \frac{1}{n}u, v\rangle$, so $\frac{1}{n}\langle u, v\rangle = \langle \frac{1}{n}u, v\rangle$. So if m, n are positive integers,

$$\langle \frac{m}{n}u, v\rangle = m\langle \frac{1}{n}u, v\rangle = \frac{m}{n}\langle u, v\rangle.$$

Using property (ii) in the definition of norm we see $||-u|| = |-1|||u|| = ||u||$. Then using the definition of $\langle\ ,\ \rangle$, we have

$$\langle -u, v\rangle = \frac{||-u+v||^2 - ||-u-v||^2|}{4} = \frac{||-(u-v)||^2 - ||-(u+v)||^2}{4}$$

$$= \frac{-||u+v||^2 + ||u-v||^2}{4} = -\langle u, v\rangle.$$

This shows that if r is any rational number, then

$$\langle ru, v\rangle = r\langle u, v\rangle.$$

Now suppose that $\lambda \in \mathcal{R}$ is any real number (with special interest in the case where λ is not rational). There must be a sequence $\{r_n\}_{n=1}^{\infty}$ of rational numbers for which $\lim_{n\to\infty} r_n = \lambda$. Thus

$$\lambda\langle u, v\rangle = \lim_{n\to\infty} r_n\langle u, v\rangle = \lim_{n\to\infty}\langle r_n u, v\rangle =$$

$$= \lim_{n\to\infty} \frac{||r_n u + v||^2 - ||r_n u - v||^2}{4}.$$

We claim that $\lim_{n\to\infty}||r_n u + v||^2 = ||\lambda u + v||^2$ and $\lim_{n\to\infty}||r_n u - v||^2 = ||\lambda u - v||^2$. Once we have shown this, we will have

$$\lambda\langle u, v\rangle = \frac{\lambda u + v||^2 - ||\lambda u - v||^2}{4} = \langle \lambda u, v\rangle,$$

completing the proof that $\langle\ ,\ \rangle$ is homogeneous in the first slot.

For $x, y \in U$, $||x|| = ||y+(x-y)|| \leq ||y||+||y-x||$, so $||x||-||y|| \leq ||x-y||$. Interchanging the roles of x and y we see that also $||y - x|| \leq ||y - x|| = ||x - y||$. Hence $|||x|| - ||y||| \leq ||y - x||$. With $x = r_n u + v$ and $y = \lambda u + v$, this latter inequality gives

$$|||r_n u + v|| - ||\lambda u + v||| \leq ||(r_n - \lambda)u|| = |r_n - \lambda| \cdot ||u||.$$

Since $r_n - \lambda \to 0$, we have $||r_n u + v|| \to ||\lambda u + v||$. Replacing v with $-v$ gives $\lim_{n\to\infty}||r_n u - v|| = ||\lambda u - v||$. So indeed $\langle\ ,\ \rangle$ is homogeneous in the first slot. Fiinally, we show that

(d) $\langle u, v\rangle = \langle v, u\rangle$.

So:

$$\langle u, v\rangle = \frac{||u + v||^2 - ||u - v||^2}{4} = \frac{||v + u||^2 - ||v - u||^2}{4} = \langle v, u\rangle.$$

This completes the proof in the case that $F = \mathcal{R}$.

Now consider the case $F = \mathcal{C}$. By the complex polarization identity (see Exercise 12) we must have

$$\langle u, v\rangle = \frac{1}{4}\sum_{n=1}^{4} i^n ||u + i^n v||^2. \tag{8.4}$$

Then putting $v = u$ we find

$$\langle u,u\rangle = \frac{1}{4}\sum_{n=1}^{4} i^n||u+i^n u||^2 = \frac{1}{4}\sum_{n=1}^{4} i^n|1+i^n|^2||u||^2$$

$$= \frac{||u||^2}{4}\left[2i-0-2i+4\right] = ||u||^2,$$

as desired. But we must still show that $\langle\ ,\ \rangle$ has the properties of an inner product.

Because $\langle u,u\ \rangle = ||u||^2$, it follows immediately from the properties of a norm that $\langle u,u\rangle \geq 0$ with equality if and only if $u=\vec{0}$.

For convenience define $\langle\ ,\ \rangle_{\mathcal{R}}$ to be the real inner product defined above, so

$$\langle u,v\rangle_{\mathcal{R}} = \frac{||u+v||^2-||u-v||^2}{4}.$$

Note that

$$\langle u,v\rangle = \langle u,v\rangle_{\mathcal{R}} + i\langle u,iv\rangle_{\mathcal{R}}.$$

We have already proved that $\langle\ ,\ \rangle_{\mathcal{R}}$ is additive in the first slot, which can now be used in a routine fashion to show that $\langle\ ,\ \rangle$ is also additive in the first slot. It is even easier to show that $\langle au,v\rangle = a\langle u,v\rangle$ for $a\in\mathcal{R}$ using the homogeneity of $\langle\ ,\ \rangle_{\mathcal{R}}$ in the first slot. However, we must still extend this to all complex numbers. Note that:

$$\langle iu,v\rangle =$$

$$= \frac{||iu+v||^2-||iu-v||^2+i||iu+iv||^2-i||iu-iv||^2}{4}$$

$$= \frac{||i(u+v)||^2 i-||i(u-v)||^2 i-||i(u+iv)||^2+||i(u-iv)||^2}{4}$$

$$= \frac{||u+v||^2 i-||u-v||^2 i-||u+iv||^2+||u-iv||^2}{4}$$

$$= i\langle u,v\rangle.$$

Combining this with additivity and homogeneity with respect to real numbers, we get that

$$\langle (a+bi)u,v\rangle = (a+bi)\langle u,v\rangle\ \forall a,b\in\mathcal{R}.$$

Hence $\langle\ ,\ \rangle$ is homogeneous in the first slot. The only thing remaining to show is that $\langle v, u\rangle = \overline{\langle u, v\rangle}$.

$$\begin{aligned}
\langle u, v\rangle &= \frac{||u+v||^2 - ||u-v||^2 + ||u+iv||^2 - ||u-iv||^2 i}{4} \\
&= \frac{||v+u||^2 - ||v-u||^2 + ||i(v-ui)||^2 - ||(-i)(v+ui)||^2 i}{4} \\
&= \overline{\frac{||v+u||^2 - ||v-u||^2 + ||v+iu||^2 i - ||v-iu||^2 i}{4}} \\
&= \overline{\langle v, u\rangle}.
\end{aligned}$$

□

Now suppose that F is a subfield of $\mathcal{C}$ and that $V = F^n$. There are three norms on V that are most commonly used in applications.

Definition For vectors $x = (x_1, \ldots, x_n)^T \in V$, the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$, called the 1-*norm*, 2-*norm*, and ∞-*norm*, respectively, are defined as:

$$\begin{aligned}
\|x\|_1 &= |x_1| + |x_2| + \cdots + |x_n|; \\
\|x\|_2 &= (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{1/2}; \\
\|x\|_\infty &= \max\{|x_1|, |x_2|, \ldots, |x_n|\}.
\end{aligned} \tag{8.5}$$

Put $x = (1, 1)^T \in F^2$ and $y = (-1, 1)^T \in F^2$. Using these vectors x and y it is routine to show that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ do not satisfy the parallelogram equality, hence must not be derived from an inner product in the usual way. On the other hand, all three norms are equivalent in a sense that we are about to make clear. First we pause to notice that the so-called norms really are norms. The only step that is challenging is the triangle inequality for the 2-norm, and we proved this earlier. The details for the other two norms are left to the reader.

Definition Let $\|\cdot\|$ be a norm on V. A sequence $\{v_i\}_{i=1}^\infty$ of vectors is said to *converge to the vector* v_∞ provided the sequence $\{\|v_i - v_\infty\|\}$ of real numbers converges to 0.

With this definition we can now talk of a sequence of vectors in F^n converging by using norms. But which norm should we use? What we mean by saying that all three norms are equivalent is that a sequence of vectors

converges to a vector v using one of the norms if and only if it converges to the same vector v using either of the other norms.

Theorem 8.1.10. *The 1-norm, 2-norm and ∞-norm on F^n are all equivalent in the sense that*

(a) If a sequence $\{x_i\}$ of vectors converges to x_∞ as determined in one *of the norms, then it converges to x_∞ in* all three *norms, and for fixed index j, the entries $(x_i)_j$ converge to the entry $(x_\infty)_j$. This is an easy consequence of*

(b) $\|x\|_1 \leq \|x\|_2\sqrt{n} \leq n\|x\|_\infty \leq n\|x\|_1$.

Proof. If $u = (u_1, \ldots, u_n)$, put $x = (|u_1|, \ldots, |u_n|)^T$ and $y = (1, 1, \ldots, 1)^T$. Then by the Cauchy-Schwarz inequality applied to x and y, we have

$$|\langle x, y\rangle| = \sum |u_i| = \|u\|_1 \leq \|x\|_2\|y\|_2 = \sqrt{\sum |u_i|^2} \cdot \sqrt{n} = \|u\|_2\sqrt{n}.$$

So $\|x\|_1 \leq \|x\|_2\sqrt{n}$ for all $x \in F^n$.

Next, $\|x\|_2^2 = x_1^2 + \cdots + x_n^2 \leq n(\max\{|x_i|\})^2 = n \cdot \|x\|_\infty^2$, implying $\|x\|_2^2 \leq \sqrt{n}\|x\|_\infty$. This proves the first and second inequalities in (b), and the third is quite obvious. □

In fact, any two norms on a finite dimensional vector space over F are equivalent (in the sense given above), but we do not need this result.

8.2 Orthonormal Bases

Theorem 8.2.1. *Let $L = (v_1, \ldots, v_m)$ be an orthonormal list of vectors in V and put $W = span(L)$. If $w = \sum_{i=0}^m a_i v_i$ is an arbitrary element of W, then*

(i) $a_i = \langle w, v_i\rangle$, and

(ii) $||w||^2 = \sum_{i=1}^m |a_i|^2 = \sum_{i=1}^m |\langle w, v_i\rangle|^2$.

Proof. If $w = \sum_{i=1}^m a_i v_i$, compute $\langle w, v_j\rangle = \langle \sum_{i=1}^m a_i v_i, v_j\rangle = a_j$. Then apply the Pythagorean Theorem. □

The preceding result shows that an orthonormal basis can be extremely handy. The next result gives the **Gram-Schmidt** algorithm for replacing a linearly independent list with an orthonormal one having the same span as the original.

Theorem 8.2.2. *(*Gram-Schmidt method*) If $L = (v_1, \ldots, v_m)$ is a linearly independent list of vectors in V, then there is an orthonormal list $\mathcal{B} = (e_1, \ldots, e_m)$ such that span$(e_1, \ldots, e_j)$ = span$(v_1, \ldots, v_j)$ for each $j = 1, 2, \ldots, m$.*

Proof. Start by putting $e_1 = v_1/||v_1||$, so e_1 has norm 1 and spans the same space as does v_1.

We construct the remaining vectors $e_2, \ldots, e_m$ inductively. Suppose that $e_1, \ldots, e_k$ have been determined so that $(e_1, \ldots, e_k)$ is orthonormal and span$(e_1, \ldots, e_j)$ = span$(v_1, \ldots, v_j)$ for each $j = 1, 2, \ldots, k$. We then construct e_{k+1} as follows. Put $e'_{k+1} = v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i$. Check that e'_{k+1} is orthogonal to each of $e_1, \ldots, e_k$. Then put $e_{k+1} = e'_{k+1}/||e'_{k+1}||$. □

At this point we need to assume that V is finite dimensional.

Corollary 8.2.3. *Let V be a finite dimensional inner product space.*

(i) V has an orthonormal basis.

(ii) If L is any orthonormal set in V it can be completed to an orthonormal basis of V.

Proof. We know that any independent set can be completed to a basis to which we can then apply the Gran-Schmidt algorithm. □

Lemma 8.2.4. *If $T \in \mathcal{L}(V)$ has an upper triangular matrix with respect to some basis of V, then it has an upper triangular matrix with respect to some orthonormal basis of V.*

Proof. Suppose that $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis such that $[T]_{\mathcal{B}}$ is upper triangular. Basically this just means that for each $j = 1, 2, \ldots, n$ the subspace span$(v_1, \ldots, v_j)$ is T-invariant. Use the Gram-Schmidt algorithm to construct an orthonormal basis $\mathcal{S} = (e_1, \ldots, e_n)$ for V such that span$(v_1, \ldots, v_j)$ = span$(e_1, \ldots, e_j)$ for each $j = 1, 2, \ldots, n$. Hence for each j, span$(e_1, \ldots, e_j)$ is T-invariant, so that $[T]_{\mathcal{S}}$ is upper triangular. □

Using the fact that $\mathcal{C}$ is algebraically closed we showed that if V is a finite dimensional complex vector space, then there is a basis $\mathcal{B}$ for V such that $[T]_{\mathcal{B}}$ is upper triangular. Hence we have the following result which is sometimes called **Schur's Theorem**.

Corollary 8.2.5. *(Schur's Theorem) If $T \in \mathcal{L}(V)$ where V is a finite dimensional complex vector space, then there is an orthonormal basis for V with respect to which T has an upper triangular matrix.*

We now apply the preceding results to the case where $V = F^n$. First we introduce a little more language. If P is an invertible matrix for which $P^{-1} = P^*$, where P^* is the conjugate transpose of P, then P is said to be *unitary*. If P is real and unitary (so $P^{-1} = P^T$), we say P is an *orthogonal* matrix. Let A be an $n \times n$ matrix over $\mathcal{C}$. View $\mathcal{C}^n$ as an inner product space with the usual inner product. Let $\mathcal{S}$ be the standard ordered (orthonormal) basis. Define $T_A \in \mathcal{L}(\mathcal{C}^n)$ by $T_A(\vec{x}) = A\vec{x}$. We know that $[T_A]_{\mathcal{S}} = A$. Let $\mathcal{B} = (v_1, \ldots, v_n)$ be an orthonormal basis with respect to which T_A has an upper triangular matrix. Let P be the matrix whose jth column is $v_j = [v_j]_{\mathcal{S}}$. Then $[T_A]_{\mathcal{B}} = P^{-1}AP$ by Theorem 4.7.2. Since $\mathcal{B}$ is orthonormal it is easy to check that P is a unitary matrix. We have proved the following.

Corollary 8.2.6. *If A is an $n \times n$ complex matrix, there is a unitary matrix P such that $P^{-1}AP$ is upper triangular.*

8.3 Orthogonal Projection and Minimization

Let V be a vector space over the field F, F a subfield of $\mathcal{C}$, and let $\langle\ ,\ \rangle$ be an inner product on V. Let W be a subspace of V, and put

$$W^{\perp} = \{v \in V : \langle w, v\rangle = 0 \text{ for all } w \in W\}.$$

Obs. 8.3.1. *$W^{\perp}$ is a subspace of V and $W \cap W^{\perp} = \{\overline{0}\}$. Hence $W + W^{\perp} = W \oplus W^{\perp}$.*

Proof. Easy exercise. □

We do not know if each $v \in V$ has a representation in the form $v = w + w'$ with $w \in W$ and $w' \in W^{\perp}$, but at least we know from the preceding observation that it has at most one. There is one case where we know that $V = W \oplus W^{\perp}$.

Theorem 8.3.2. *If W is finite dimensional, then $V = W \oplus W^{\perp}$.*

Proof. Suppose W is finite dimensional. Then using the Gram-Schmidt process, for example, we can find an orthonormal basis $\mathcal{D} = (v_1, \ldots, v_m)$ of W. For arbitrary $v \in V$, put $a_i = \langle v, v_i\rangle$. Then we know that

$$w = \sum_{i=1}^{m} a_i v_i$$

is in W, and we write

$$v = w + w' = \sum_{i=1}^{m} a_i v_i + (v - \sum_{i=1}^{m} a_i v_i).$$

We show that $v - \sum_{i=1}^{m} a_i v_i \in W^\perp$.

So let $u \in W$, say $u = \sum_{j=1}^{m} b_j v_j$. Since $\mathcal{D}$ is orthonormal,

$$\langle v - \sum_{1}^{m} a_i v_i, u\rangle = \langle v - \sum_{i=1}^{m} a_i v_i, \sum_{j=1}^{m} b_j v_j\rangle$$

$$= \sum_{j=1}^{m} \overline{b}_j \langle v, v_j\rangle - \sum_{i,j=1}^{m} \overline{b}_j a_i \langle v_i, v_j\rangle$$

$$= \sum_{j=1}^{m} \overline{b}_j a_j - \sum_{j=1}^{m} \overline{b}_j a_j \langle v_j, v_j\rangle = 0.$$

So with $w = \sum_{i=1}^{m} a_i v_i$ and $w' = v - w$, $v = w + w'$ is the unique way to write v as the sum of an element of W plus an element of $W^\perp$. □

Now suppose $V = W \oplus W^\perp$, (which we have just seen is the case if W is finite dimensional). A linear transformation $P : V \to V$ is said to be an *orthogonal projection* of V onto W provided the following hold:

(i) $P(w) = w$ for all $w \in W$,

(ii) $P(w') = \overline{0}$ for all $w' \in W^\perp$.

So let P be an orthogonal projection onto W by this definition. Let $v = w + w'$ be any element of V with $w \in W$, $w' \in W^\perp$. Then $P(v) = P(w + w') = P(w) + P(w') = w + \overline{0} = w \in W$. So $P(v)$ is a uniquely defined element of W for all v.

Conversely, still under the hypothesis that $V = W \oplus W^\perp$, define $P' : V \to V$ as follows. For $v \in V$, write $v = w + w'$, $w \in W$, $w' \in W^\perp$ (uniquely!), and put $P'(v) = w$. It is an easy exercise to show that P' really is linear, $P'(w) = w$ for all $w \in W$, and $P'(w') = P'(\overline{0} + w') = \overline{0}$ for $w' \in W^\perp$. So $P' : v = w + w' \mapsto w$ is really *the unique orthogonal projection* of V onto W. Moreover,

Obs. 8.3.3. *$P^2 = P$; $P(v) = v$ if and only if $v \in W$; $P(v) = \overline{0}$ if and only if $v \in W^\perp$.*

Obs. 8.3.4. *$I - P$ is the unique projection of V onto $W^\perp$.*

Both Obs. 8.3.3 and 8.3.4 are fairly easy to prove, and their proofs are worthwhile exercises.

Obs. 8.3.5. $W \subseteq W^{\perp\perp}$.

Proof. $W^\perp$ consists of all vectors in V orthogonal to every vector of W. So in particular, every vector of W is orthogonal to every vector of $W^\perp$, i.e., each vector of W is in $(W^\perp)^\perp$. But in general we do not know if there could be some vector outside W that is in $W^{\perp\perp}$. □

Theorem 8.3.6. *If $V = W \oplus W^\perp$, then $W = (W^\perp)^\perp$.*

Proof. By Obs. 8.3.5, we must show that $W^{\perp\perp} \subseteq W$. Since $V = W \oplus W^\perp$, we know there is a unique orthogonal projection P of V onto W, and for each $v \in V$, $v - P(v) \in W^\perp$. Keep in mind that $W \subseteq (W^\perp)^\perp$ and $P(v) \in W \subseteq (W^\perp)^\perp$, so $\langle v - P(v), P(v)\rangle = \overline{0}$. It follows that for $v \in W^{\perp\perp}$ we have

$$||v-P(v)||^2 = \langle v-P(v), v-P(v)\rangle = \langle v, v-P(v)\rangle - \langle P(v), v-P(v)\rangle = \overline{0}-\overline{0} = \overline{0}$$

by the comments just above. Hence $v = P(v)$, which implies that $v \in W$, i.e., $W^{\perp\perp} \subseteq W$. Hence $W^{\perp\perp} = W$. □

Note: If $V = W \oplus W^\perp$, then $(W^\perp)^\perp = W$, so also $V = W^\perp \oplus (W^\perp)^\perp$, and $(W^\perp)^{\perp\perp} = W^\perp$.

Given a finite dimensional subspace U of the inner product space V and a point $v \in V$, we want to find a point $u \in U$ closest to v in the sense that $||v - u||$ is as small as possible. To do this we first construct an orthonormal basis $\mathcal{B} = (e_1, \ldots, e_m)$ of U. The unique orthogonal projection of V onto U is given by

$$P_U(v) = \sum_{i=0}^{m} \langle v, e_i\rangle e_i.$$

We show that $P_U(v)$ is the unique $u \in U$ closest to v.

Theorem 8.3.7. *Suppose U is a finite dimensional subspace of the inner product space V and $v \in V$. Then*

$$||v - P_U(v)|| \leq ||v - u|| \quad \forall u \in U.$$

Furthermore, if $u \in U$ and equality holds, then $u = P_U(v)$.

Proof. Suppose $u \in U$. Then $v - P_U(v) \in U^\perp$ and $P_U(v) - u \in U$, so we may use the Pythagorean Theorem in the following:

$$\begin{aligned} ||v - P_U(v)||^2 &\leq ||v - P_U(v)||^2 + ||P_U(v) - u||^2 \\ &= ||(v - P_U(v)) + (P_U(v) - u)||^2 \\ &= ||v - u||^2, \end{aligned} \tag{8.6}$$

where taking square roots gives the desired inequality. Also, the inequality of the theorem is an equality if and only if the inequality in Eq. 8.6 is an equality, which is if and only if $u = P_U(v)$. □

Example 8.3.8. *There are many applications of the above theorem. Here is one example. Let $V = \mathcal{R}[x]$, and let U be the subspace consisting of all polynomials $f(x)$ with degree less than 4 and satisfying $f(0) = f'(0) = 0$. Find a polynomial $p(x) \in U$ such that $\int_0^1 |2 + 3x - p(x)|^2 dx$ is as small as possible.*

Solution: Define an inner product on $\mathcal{R}[x]$ by $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$. Put $g(x) = 2 + 3x$, and note that $U = \{p(x) = a_2x^2 + a_3x^3 : a_2, a_3 \in \mathcal{R}\}$. We need an orthonormal basis of U. So start with the basis $\mathcal{B} = (x^2, x^3)$ and apply the Gram-Schmidt algorithm. We want to put $e_1 = \frac{x^2}{||x^2||}$. First compute $||x^2||^2 = \int_0^1 x^4 dx = \frac{1}{5}$, so

$$e_1 = \sqrt{5}\, x^2. \tag{8.7}$$

Next we want to put

$$e_2 = \frac{x^3 - \langle x^3, e_1\rangle e_1}{||x^3 - \langle x^3, e_1\rangle e_1||}.$$

Here $\langle x^3, e_1\rangle = \sqrt{5}\int_0^1 x^5 dx = \frac{\sqrt{5}}{6}$. Then $x^3 - \langle x^3, e_1\rangle e_1 = x^3 - \frac{5}{6}x^2$.

Now $||x^3 - \frac{5}{6}x^2||^2 = \int_0^1 (x^3 - \frac{5}{6}x^2)^2 dx = \frac{1}{7\cdot 36}$. Hence

$$e_2 = 6\sqrt{7}(x^3 - \frac{5}{6}x^2) = \sqrt{7}(6x^2 - 5x^2). \tag{8.8}$$

Then the point $p \in U$ closest to $g = 2 + 3x$ is

$$\begin{aligned} p &= \langle g, e_1\rangle e_1 + \langle g, e_2\rangle e_2 \\ &= \left[\sqrt{5}\int_0^1 (2+3x)x^2 dx\right]\sqrt{5}x^2 + \qquad (8.9) \\ &+ \left[\sqrt{7}\int_0^1 (2+3x)(6x^3-5x^2)dx\right]\sqrt{7}(6x^3-5x^2), \end{aligned}$$

(8.10)

so that after a bit more computation we have

$$p = 24x^2 - \frac{203}{10}x^3. \qquad (8.11)$$

8.4 Linear Functionals and Adjoints

Recall that if V is a vector space over any field F, then a linear map from V to F (viewed as a vector space over F) is called a *linear functional*. The set V^* of all linear functionals on V is called the *dual space* of V. When V is an inner product space the linear functionals on V have a particularly nice form. Fix $v \in V$. Then define $\phi_v : V \to F : u \mapsto \langle u, v\rangle$. It is easy to see that ϕ_v is a linear functional. If V is finite dimensional then every linear functional on V arises this way.

Theorem 8.4.1. *Let $\phi \in V^*$. Then there is a unique vector $v \in V$ such that*

$$\phi(u) = \langle u, \rangle v$$

for every $u \in V$.

Proof. Let $(e_1, \ldots, e_n)$ be an orthonormal basis of V. Then for any $u \in V$, we have

$$u = \sum_{i=1}^{n} \langle u, e_i\rangle e_i.$$

Hence

$$\begin{aligned} \phi(u) &= \phi(\sum_{i=1}^{n}\langle u, e_i\rangle e_i) = \sum_{i=1}^{n}\langle u, e_i\rangle\phi(e_i) \\ &= \sum_{i=1}^{n}\langle u, \overline{\phi(e_i)}\rangle e_i. \end{aligned}$$

(8.12)

So if we put $v = \sum_{i=1}^{n} \overline{\phi(e_i)} e_i$, we have $\phi(u) = \langle u, v \rangle$ for every $u \in V$. This shows the existence of the desired v. For the uniqueness, suppose that

$$\phi(u) = \langle u, v_1 \rangle = \langle u, v_2 \rangle \quad \forall u \in V.$$

Then $0 = \langle u, v_1 - v_2 \rangle$ for all $u \in V$, forcing $v_1 = v_2$. □

Now let V and W both be inner product spaces over F. Let $T \in \mathcal{L}(V, W)$ and fix $w \in W$. Define $\phi : V \to F$ by $\phi(v) = \langle T(v), w \rangle$. First, it is easy to check that $\phi \in V^*$. Second, by the previous theorem there is a unique vector (that we now denote by $T^*(w)$) for which

$$\phi(v) = \langle T(v), w \rangle = \langle v, T^*(w) \rangle \quad \forall v \in V.$$

It is clear that T^* is some kind of map from W to V. In fact, it is routine to check that T^* is linear, i.e., $T^* \in \mathcal{L}(W, V)$.

Theorem 8.4.2. *Let V and W be inner product spaces and suppose that $S, T \in \mathcal{L}(V, W)$ are both such that S^* and T^* exist. Then*
(i) $(S + T)^* = S^* + T^*$.
(ii) $(aT)^* = \overline{a} T^*$.
(iii) $(T^*)^* = T$.
(iv) $I^* = I$.
(v) $(ST)^* = T^* S^*$.

Proof. The routine proofs are left to the reader. □

Theorem 8.4.3. *Let V and W be finite dimensional inner product spaces over F. Suppose $T\mathcal{L}(V, W)$. Then*
(i) $null(T^*) = (Im(T))^{\perp}$;
(ii) $Im(T^*) = (null(T))^{\perp}$;
(iii) $null(T) = (Im(T^*))^{\perp}$;
(iv) $Im(T) = (null(T^*))^{\perp}$.

Proof. For (i), $w \in \text{null}(T^*)$ iff $T^*(w) = \vec{0}$ iff $\langle v, T^*(w) \rangle = 0 \quad \forall v \in V$ iff $\langle T(v), w \rangle = 0 \quad \forall v \in V$ iff $w \in (\text{Im}(T))^{\perp}$. This proves (i). By taking orthogonal complements of both sides of an equality, or by replacing an operator with its adjoint, the other three equalities are easily established. □

Theorem 8.4.4. *Let V be a finite dimensional inner product space over F and let $\mathcal{B} = (e_1, \ldots, e_n)$ be an orthonormal basis for V. Let $T \in \mathcal{L}(V)$ and let $A = [T]_{\mathcal{B}}$. Then $A_{ij} = \langle T(e_j), e_i \rangle$.*

Proof. The matrix A is defined by

$$T(e_j) = \sum_{i=1}^{n} A_{ij} e_i.$$

Since $\mathcal{B}$ is orthonormal, we also have

$$T(e_j) = \sum_{i=1}^{n} \langle T(e_j), e_i \rangle e_i.$$

Hence $A_{ij} = \langle T(e_j), e_i \rangle$. □

Corollary 8.4.5. *Let V be a finite dimensional inner product space over F and let $\mathcal{B} = (e_1, \ldots, e_n)$ be an orthonormal basis for V. Let $T \in \mathcal{L}(V)$ and let $A = [T]_{\mathcal{B}}$. Then $[T^*]_{\mathcal{B}} = A^*$, where A^* is the conjugate transpose of A.*

Proof. According to Theorem 8.4.4, $A_{ij} = \langle T(e_j), e_i \rangle$, and if $B = [T^*]_{\mathcal{B}}$, then $B_{ij} = \langle T^*(e_j), e_i \rangle = \overline{\langle e_i, T^*(e_j) \rangle} = \overline{\langle T(e_i), e_j \rangle} = \overline{A_{ji}}$. □

8.5 The Rayleigh Principle*

All matrices in this note are over the field $\mathcal{C}$ of complex numbers, and for any matrix B, B^* denotes the conjugate transpose of B. Also, $\langle\ ,\ \rangle$ denotes the standard inner product on $\mathcal{C}^n$ given by: $\langle x, y \rangle = x^T \overline{y} = y^* x$. Let A be $n \times n$ hermitian, so A has real eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ with an orthonormal set $\{v_1, \ldots, v_n\}$ of associated eigenvectors: $Av_j = \lambda_j v_j$, $\langle v_j, v_i \rangle = v_j^T \overline{v_i} = v_i^* v_j = \delta_{ij}$.

Let $Q = (v_1, \ldots, v_n)$ be the matrix whose jth column is v_j. Then $Q^*Q = I_n$ and $Q^*AQ = Q^*(\lambda_1 v_1, \ldots, \lambda_n v_n) = Q^*Q\Lambda = \Lambda = diag(\lambda_1, \ldots, \lambda_n)$, and Q is unitary ($Q^* = Q^{-1}$).

For $\overline{0} \neq x \in \mathcal{C}^n$ define the *Rayleigh Quotient* $\rho_A(x)$ for A by

$$\rho_A(x) = \frac{\langle Ax, x \rangle}{\langle x, x \rangle} = \frac{x^* Ax}{||x||^2}. \tag{8.13}$$

Put $\mathcal{O} = \{x \in \mathcal{C}^n : \langle x, x \rangle = 1\}$, and note that for $0 \neq k \in \mathcal{C}, \overline{0} \neq x \in \mathcal{C}^n$,

$$\rho_A(kx) = \rho_A(x). \tag{8.14}$$

Hence

$$\{\rho_A(x) : x \neq 0\} = \{\rho_A(x) : x \in \mathcal{O}\} = \{x^*Ax : x \in \mathcal{O}\}. \tag{8.15}$$

The set $W(A) = \{\rho_A(x) : x \in \mathcal{O}\}$ is called the *numerical range* of A. Observe that if $x = Qy$, then $x \in \mathcal{O}$ iff $y \in \mathcal{O}$. Since $W(A)$ is the continuous image of a compact connected set, it must be a closed bounded interval with a maximum M and a minimum m. Since $Q : \mathcal{C}^n \to \mathcal{C}^n : x \mapsto Qx$ is nonsingular, Q maps $\mathcal{O}$ to $\mathcal{O}$ in a one-to-one and onto manner. Hence

$$M = max_{x\in\mathcal{O}}\{x^*Ax\} = max_{y\in\mathcal{O}}\{(Qy)^*A(Qy) = y^*Q^*AQy\} =$$

$$= max_{y\in\mathcal{O}}\{y^*\Lambda y\} = max_{y\in\mathcal{O}}\{\sum_{j=1}^{n} \lambda_j|y_j|^2\},$$

where $y = (y_1, y_2, \ldots, y_n)^T \in \mathcal{C}^n$, $\sum |y_i|^2 = 1$.

Similarly,

$$m = min_{x\in\mathcal{O}}\{\rho_A(x)\} = min_{y\in\mathcal{O}}\{\sum_{j=1}^{n} \lambda_j|y_j|^2\}.$$

By the ordering of the eigenvalues, for $y \in \mathcal{O}$ we have

$$\lambda_1 = \lambda_1 \sum |y_j|^2 \leq \sum_{j=1}^{n} \lambda_j|y_j|^2 = \rho_\Lambda(y) \leq \lambda_n \sum_{j=1}^{n} |y_j|^2 = \lambda_n. \tag{8.16}$$

Furthermore, with $y = (1, 0, \ldots, 0)^* \in \mathcal{O}$ and $z = (0, \ldots, 0, 1)^* \in \mathcal{O}$, we have $\rho_\Lambda(y) = \lambda_1$ and $\rho_\Lambda(z) = \lambda_n$. Hence we have almost proved the following first approximation to the *Rayleigh Principle.*

Theorem 8.5.1. *Let A be an $n \times n$ hermitian matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then for any nonzero $x \in \mathcal{O}^n$,*

$$\lambda_1 \leq \rho_A(x) \leq \lambda_n, \text{ and} \tag{8.17}$$

$$\lambda_1 = min_{x\in\mathcal{O}}\rho_A(x); \quad \lambda_n = max_{x\in\mathcal{O}}\rho_A(x). \tag{8.18}$$

$$\text{If } \overline{0} \neq x \in \mathcal{C}^n \text{ satisfies } \rho_A(x) = \lambda_i \text{ for either } i = 1 \text{ or } i = n, \tag{8.19}$$

then x is an eigenvector of A belonging to the eigenvalue λ_i.

Proof. Clearly Eqs. 8.17 and 8.18 are already proved. So consider Eq. 8.19.

Without loss of generality we may assume $x \in \mathcal{O}$. Suppose $x = \sum_{j=1}^n c_j v_j$, so that $\rho_A(x) = x^* A x = \left(\sum_{j=1}^n \overline{c}_j v_j^*\right)\left(\sum_{j=1}^n c_j \lambda_j v_j\right) = \sum_{j=1}^n \lambda_j |c_j|^2$.

Clearly $\lambda_1 = \lambda_1 \sum_{j=1}^n |c_j|^2 \leq \sum_{j=1}^n \lambda_j |c_j|^2$ with equality iff $\lambda_j = \lambda_1$ whenever $c_j \neq 0$. Hence $\rho_A(x) = \lambda_1$ iff x belongs to the eigenspace associated with λ_1. The argument for λ_n is similar. □

Recall that $Q = (v_1, \ldots, v_n)$, and note that if $x = Qy$, so $y = Q^* x$, then $x = v_i = Qy$ iff $y = Q^* v_i = \overline{e}_i$. So with the notation $x = Qy$, $y = (y_1, \ldots, y_n)^T$, we have

$$\langle x, v_i \rangle = x^* v_i = x^* Q Q^* v_i = y^* \overline{e}_i = \overline{y}_i.$$

Hence $x \perp v_i$ iff $y = Q^* x$ satisfies $y_i = 0$.

Def. $T_j = \{x \neq 0 : \langle x, v_k \rangle = 0 \text{ for } k = 1, \ldots, j\} = \{v_1, \ldots, v_j\}^\perp \setminus \{\overline{0}\}$.

Theorem 8.5.2. *$\rho_A(x) \geq \lambda_{j+1}$ for all $x \in T_j$, and $\rho_A(x) = \lambda_{j+1}$ for some $x \in T_j$ iff x is an eigenvector associated with λ_{j+1}. Thus*

$$\lambda_{j+1} = min_{x \in T_j}\{\rho_A(x)\} = \rho_A(v_{j+1}).$$

Proof. $\overline{0} \neq x \in T_j$ iff $x = Qy$ where $y = \sum_{k=j+1}^n y_k \overline{e}_k$ iff $x = \sum_{k=j+1}^n y_k v_k$. Without loss of generality we may assume $x \in \mathcal{O} \cap T_j$. Then $y = Q^* x \in \mathcal{O}$ and $x \in T_j \cap \mathcal{O}$ iff $\rho_A(x) = \sum_{k=j+1}^n \lambda_k |y_k|^2 \geq \lambda_{j+1} \sum_{k=j+1}^n |y_k^2| \geq \lambda_{j+1}$, with equality iff $y_k = 0$ whenever $\lambda_k > \lambda_{j+1}$. In particular, if $y = \overline{e}_{j+1}$, so $x = v_{j+1},\ \rho_A(x) = \lambda_{j+1}$. □

Theorem 8.5.3. *Put $S_j = \{x \neq 0 : \langle x, v_k \rangle = 0$ for $k = n, n-1, \ldots, n-(j-1)\}$. So $S_j = \{v_n, v_{n-1}, \ldots, v_{n-(j-1)}\}^\perp \setminus \{\vec{0}\}$. Then $\rho_A(x) \leq \lambda_{n-j}$ for all $x \in S_j$, and equality holds iff x is an eigenvector associated with λ_{n-j}. Thus*

$$\lambda_{n-j} = max_{x \in S_j}\{\rho_A(x)\} = \rho_A(v_{n-j}).$$

Proof. We leave the proof as an exercise for the reader. □

The *Rayleigh Principle* consists of Theorems 8.5.2 and 8.5.3.

8.6 Exercises

1. **Parallelogram Equality:** If u, v are vectors in the inner product space V, then

$$||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2).$$

   Explain why this equality should be so-named.

2. Let V and W be inner product spaces over F, and let $T \in \mathcal{L}(V,W)$. If there is an adjoint map $T^* : W \to V$ such that $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ for all $v \in V$ and all $w \in W$, show that $T^* \in \mathcal{L}(W,V)$.

3. Let $T : V \to W$ be a linear transformation whose adjoint $T^* : W \to V$ with respect to $\langle\ ,\ \rangle_V$, $\langle\ ,\ \rangle_W$ does exist. So for all $v \in V$, $w \in W$, $\langle T(v), w\rangle_W = \langle v, T^*(w)\rangle_V$. Put $N = null(T)$ and $R^* = Im(T^*)$.

   (a) Show that $(R^*)^\perp = N$.

   (b) Suppose that R^* is finite dimensional and show that $N^\perp = R^*$.

4. Prove Theorem 8.4.2.

5. Prove Theorem 8.5.3.

6. On an in-class linear algebra exam, a student was asked to give (for ten points) the definition of "real symmetric matrix." He couldn't remember the definition and offered the following alternative: A real, $n \times n$ matrix A is symmetric if and only if $A^2 = AA^T$. When he received no credit for this answer, he went to see the instructor to find out what was wrong with his definition. By that time he had looked up the definition and tried to see if he could prove that his definition was equivalent to the standard one: A is symmetric if and only if $A = A^T$. He had a proof that worked for 2×2 matrices and thought he could prove it for 3×3 matrices. The instructor said this was not good enough. However, the instructor would give the student 5 points for a counterexample, or the full ten points if he could prove that the two definitions were equivalent for all n. Your problem is to determine whether or not the student should have been able to earn ten points or merely five points. And you might consider the complex case also: show that $A^2 = AA^*$ if and only $A = A^*$, or find a counterexample.

7. Suppose V and W are finite dimensional inner product spaces with orthonormal bases $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively. Let $T \in \mathcal{L}(V,W)$, so we know that $T^* \in \mathcal{L}(W,V)$ exists and is unique. Prove that $[T^*]_{\mathcal{B}_1,\mathcal{B}_2}$ is the conjugate transpose of $[T]_{\mathcal{B}_2,\mathcal{B}_1}$.

8. Suppose that V is an inner product space over F. Let $u, v \in V$. Prove that $\langle u, v\rangle = 0$ iff $||u|| \leq ||u + av|| \ \ \forall a \in F$.

   (Hint: If $u \perp v$, use the Pythagorean theorem on $u + av$. For the converse, square the assumed inequality and then show that $-2\text{Re}(\overline{a}\langle u, v\rangle) \leq |a|^2||v||$ for all $a \in F$. Then put $a = -1/||v||$ in the case $v \neq \vec{0}$.)

9. For arbitrary real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, show that

$$\left(\sum_{j=1}^{n} a_j b_j\right)^2 \leq \left(\sum_{j=1}^{n} j a_j^2\right)\left(\sum_{j=1}^{n} b_j^2/j\right).$$

10. Let V be a real inner product space. Show that

$$\langle u, v\rangle = \frac{1}{4}||u + v||^2 - \frac{1}{4}||u - v||^2. \tag{8.20}$$

    (This equality is the *real polarization identity.*)

11. Let V be a complex inner product space. Show that

$$\langle u, v\rangle = \text{Re}(\langle u, v\rangle) + i\ \text{Re}(\langle u, iv\rangle). \tag{8.21}$$

12. Let V be a complex inner product space. Show that

$$\langle u, v\rangle = \frac{1}{4}\sum_{n=1}^{4} i^n ||u + i^n v||^2. \tag{8.22}$$

    (This equality is the *complex polarization identity.*)

    For the next two exercises let V be a finite dimensional inner product space, and let $P \in \mathcal{L}(V)$ satisfy $P^2 = P$.

13. Show that P is an orthogonal projection if and only if $\text{null}(P) \subseteq (\text{Im}(P))^{\perp}$.

14. Show that P is an orthogonal projection if and only if $||P(v)|| \leq ||v||$ for all $v \in V$. (Hint: You will probably need to use Exercises 8 and 13.)

15. Fix a vector $v \in V$ and define $T \in V^*$ by $T(u) = \langle u, v \rangle$. For $a \in F$ determine the formula for $T^*(a)$. Here we us the standard inner product on F given by $\langle a, b \rangle = a\overline{b}$ for $a, b \in F$.

16. Let $T \in \mathcal{L}(V, W)$. Prove that

    (a) T is injective if and only if T^* is surjective;

    (b) T is surjective if and only if T^* is injective.

17. If $T \in \mathcal{L}(V)$ and U is a T-invariant subspace of V, then $U^\perp$ is T^*-invariant.

18. Let V be a vector space with norm $\|\cdot\|$. Prove that

$$| \, \|u\| - \|v\| \, | \leq \|u - v\|.$$

19. Use exercise 18 to show that if $\{v_i\}_{i=1}^\infty$ converges to v in V, then $\{\|u_i\|\}_{i=1}^\infty$ converges to $\|v\|$. Give an example in $\mathcal{R}^2$ to show that the norms may converge without the vectors converging.

Chapter 9

Operators on Inner Product Spaces

Throughout this chapter V will denote an inner product space over F. To make life simpler we also assume that V is finite dimensional, so operators on V always have adjoints, etc.

9.1 Self-Adjoint Operators

An operator $T \in \mathcal{L}(V)$ is called *Hermitian* or *self-adjoint* provided $T = T^*$.

Theorem 9.1.1. *Let $F = \mathcal{C}$, i.e., V is a complex inner product space, and suppose $T \in \mathcal{L}(V)$. Then*

(a) If $\langle T(v), v\rangle = 0 \ \ \forall v \in V$, then $T = 0$.

(b) If T is self-adjoint, then each eigenvalue of T is real.

(c) T is self-adjoint if and only if $\langle T(v), v\rangle \in \mathcal{R} \ \ \forall v \in V$.

Proof. It is routine to show that for all $u, w \in V$,

$$\begin{aligned}\langle T(u), w\rangle &= \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4} + \\ &+ \frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4} i.\end{aligned}$$

Note that each term on the right hand side is of the form $\langle T(v), v\rangle$ for an appropriate $v \in V$, so by hypothesis $\langle T(u), w\rangle = 0$ for all $u, w \in V$. Put $w = T(u)$ to conclude that $T = 0$. This proves part (a).

For part (b), suppose that $T = T^*$, and let v be a nonzero vector in V such that $T(v) = \lambda v$ for some complex number λ. Then

$$\lambda\langle v, v\rangle = \langle T(v), v\rangle = \langle v, T(v)\rangle = \overline{\lambda}\langle v, v\rangle.$$

Hence $\lambda = \overline{\lambda}$, implying $\lambda \in \mathcal{R}$, proving part (b).

For part (c) first suppose that $\langle T(v), v\rangle \in \mathcal{R}$ for all $v \in V$, and let $v \in V$. Then

$$\begin{aligned} \langle T(v), v\rangle - \overline{\langle T(v), v\rangle} &= \langle T(v), v\rangle - \langle v, T(v)\rangle \\ &= \langle T(v), v\rangle - \langle T^*(v), v\rangle \\ &= \langle (T - T^*)(v), v\rangle. \end{aligned}$$

If $\langle T(v), v\rangle \in \mathcal{R}$ for every $v \in V$, then the left side of the equation equals 0, so $\langle (T - T^*)(v), v\rangle = 0$ for each $v \in V$. Hence by part (a), $T - T^* = 0$, i.e., T is self-adjoint.

Conversely, suppose T is self-adjoint.Then the right hand side of the equation above equals 0, so the left hand side must also be 0, implying $\langle T(v), v\rangle \in \mathcal{R}$ for all $v \in V$, as claimed. □

Now suppose that V is a real inner product space. Consider the operator $T \in \mathcal{L}(\mathcal{R}^2)$ defined by $T(x, y) = -y, x)$ with the standard inner product on $\mathcal{R}^2$. Then $\langle T(v), v\rangle = 0$ for all $v \in \mathcal{R}^2$ but $T \neq 0$. However, this cannot happen if T is self-adjoint.

Theorem 9.1.2. *Let T be a self-adjoint operator on the real inner product space V and suppose that $\langle T(v), v\rangle = 0$ for all $v \in V$. Then $T = 0$.*

Proof. Suppose the hypotheses of the theorem hold. It is routine to verify

$$\langle T(u), w\rangle = \frac{\langle T(u + w), u + w\rangle - \langle T(u - w), u - w\rangle}{4}$$

using $\langle T(w), u\rangle = \langle w, T(u)\rangle = \langle T(u), w\rangle$ because T is self-adjoint and V is a real vector space. Hence also $\langle T(u), w\rangle = 0$ for all $u, w \in V$. With $w = T(u)$ we see $T = 0$. □

Lemma 9.1.3. *Let F be any subfield of $\mathcal{C}$ and let $T \in \mathcal{L}(V)$ be self-adjoint. If $\alpha, \beta \in \mathcal{R}$ are such that $x^2 + \alpha x + \beta$ is irreducible over $\mathcal{R}$, i.e., $\alpha^2 < 4\beta$, then $T^2 + \alpha T + \beta I$ is invertible.*

Proof. Suppose $\alpha^2 < 4\beta$ and $\vec{0} \neq v \in V$. Then

$$\begin{aligned}
\langle (T^2 + \alpha T + \beta I)(v), v\rangle &= \langle T^2(v), v\rangle + \alpha\langle T(v), T\rangle + \beta\langle v, v\rangle \\
&= \langle T(v), T(v)\rangle + \alpha\langle T(v), v\rangle + \beta||v||^2 \\
&\geq ||T(v)||^2 - |\alpha|\ ||T(v)||\ ||v|| + \beta||v||^2 \\
&= \left(||T(v)|| - \frac{|\alpha|\ ||v||}{2}\right)^2 + \left(\beta - \frac{\alpha^2}{4}\right)||v||^2 \\
&> 0, \qquad (9.1)
\end{aligned}$$

where the first inequality holds by the Cauchy-Schwarz inequality. The last inequality implies that $(T^2+\alpha T+\beta I)(v) \neq \vec{0}$. Thus $T^2+\alpha T+\beta I$ is injective, hence it is invertible. □

We have seen that some operators on a real vector space fail to have eigenvalues, but now we show that this cannot happen with self-adjoint operators.

Lemma 9.1.4. *Let T be a self-adjoint linear operator on the real vector space V. Then T has an eigenvalue.*

Proof. Suppose $n = \dim(V)$ and choose $v \in V$ with $\vec{0} \neq v$. Then

$$(v, T(v), \ldots, T^n(v))$$

must be linearly dependent. Hence there exist real numbers $a_0, \ldots, a_n$, not all 0, such that

$$\vec{0} = a_0 v + a_1 T(v) + \cdots + a_n T^n(v).$$

Construct the polynomial $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ which can be written in factored form as

$$f(x) = c(x^2 + \alpha_1 x + \beta_1) \cdots (x^2 + \alpha_k x + \beta_k)(x - \lambda_1 \cdots (x - \lambda_m),$$

where c is a nonzero real number, each α_j, β_j and λ_j is real, each $\alpha_j^2 < 4\beta_j$, $m + k \geq 1$, and the equation holds for all real x. Then we have

$$\begin{aligned}
0 &= a_0 v + a_1 T(v) + \cdots + a_n T^n(v) \\
&= (a_0 I + a_1 T + \cdots + a_n T^n)(v) \\
&= c(T^2 + \alpha_1 T + \beta_1 I) \cdots (T^2 + \alpha_k T + \beta_k)(T - \lambda_1 I) \cdots (T - \lambda_m I)(v).
\end{aligned}$$

Each $T^2+\alpha_j T+\beta_j I$ is invertible by Lemma 9.1.3 because T is self-adjoint and each $\alpha_j^2 < 4\beta_j$. Also $c \neq 0$. Hence the equation above implies that

$$0 = (T - \lambda_1 I) \cdots (T - \lambda_m I)(v).$$

It then follows that $T - \lambda_j I$ is not injective for at least one j. This says that T has an eigenvalue. □

The next theorem is very important for operators on real inner product spaces.

Theorem 9.1.5. The Real Spectral Theorem: *Let T be an operator on the real inner product space V. Then V has an orthonormal basis consisting of eigenvectors of T if and only if T is self-adjoint.*

Proof. First suppose that V has an orthonormal basis $\mathcal{B}$ consisting of eigenvectors of T. Then $[T]_{\mathcal{B}}$ is a real diagonal matrix, so it equals its conjugate transpose, i.e., T is self-adjoint.

For the converse, suppose that T is self-adjoint. Our proof is by induction on $n = \dim(V)$. The desired result clearly holds if $n = 1$. So assume that $\dim(V) = n > 1$ and that the desired result holds on vector spaces of smaller dimension. By Lemma 9.1.4 we know that T has an eigenvalue λ with a nonzero eigenvector u, and without loss of generality we may assume that $||u|| = 1$. Let $U = \text{span}(u)$. Suppose $v \in U^\perp$, i.e. $\langle u, v\rangle = 0$. Then

$$\langle u, T(v)\rangle = \langle T(u), v\rangle = \lambda\langle u, v\rangle = 0,$$

so $T(v) \in U^\perp$ whenever $u \in U^\perp$, showing that $U^\perp$ is T-invariant. Hence the map $S = T|_{U^\perp} \in \mathcal{L}(U^\perp)$. If $v, w \in U^\perp$, then

$$\langle S(v), w\rangle = \langle T(v), w\rangle = \langle v, T(w)\rangle = \langle v, S(w)\rangle,$$

which shows that S is self-adjoint. Thus by the induction hypothesis there is an orthonormal basis of $U^\perp$ consisting of eigenvectors of S. Clearly every eigenvector of S is an eigenvector of T. Thus adjoining u to an orthonormal basis of $U^\perp$ consisting of eigenvectors of S gives an orthonormal basis of V consisting of eigenvectors of T, as desired. □

Corollary 9.1.6. *Let A be a real $n \times n$ matrix. Then there is an orthogonal matrix P such that $P^{-1}AP$ is a (necessarily real) diagonal matrix if and only if A is symmetric.*

Corollary 9.1.7. *Let A be a real symmetric matrix with distinct (necessarily real) eigenvalues $\lambda_1, \ldots, \lambda_m$. Then*

$$V = null(T - \lambda_1 I) \oplus \cdots \oplus null(T - \lambda_m I).$$

9.2 Normal Operators

Definition An operator $T \in \mathcal{L}(V)$ is called *normal* provided $TT^* = T^*T$. Clearly any self-adjoint operator is normal, but there are many normal operators in general that are not self-adjoint. For example, if A is an $n \times n$ nonzero real matrix with $A^T = -A$ (i.e., A is *skew-symmetric* , then $A \neq A^*$ but A is a *normal matrix* because $AA^* = A^*A$. It follows that if $T \in \mathcal{L}(\mathcal{R}^n)$ is the operator with $[T]_{\mathcal{S}} = A$ where $\mathcal{S}$ is the standard basis of $\mathcal{R}^n$, then T is normal but not self-adjoint.

Recall Theorem 7.1.2 (A list of nonzero eigenvectors belonging to distinct eigenvalues must be linearly independent.)

Theorem 9.2.1. *Let $T \in \mathcal{L}(V)$. Then $||T(v)|| = ||T^*(v)|| \ \ \forall v \in V$ iff T is normal.*

Proof.

$$\begin{aligned} T \text{ is normal} &\iff T^*T - TT^* = 0 \\ &\iff \langle (T^*T - TT^*)(v), v\rangle = 0 \ \forall v \in V \\ &\iff \langle T^*T(v), v\rangle = \langle TT^*(v), v\rangle \ \forall v \in V \\ &\iff ||T(v)||^2 = ||T^*(v)||^2 \ \forall v \in V. \end{aligned} \tag{9.2}$$

Since $T^*T - TT^*$ is self-adjoint, the theorem follows from Theorems 9.1.1 part (a) and 9.1.2. □

Corollary 9.2.2. *Let $T \in \mathcal{L}(V)$ be normal. Then*

(a) If $v \in V$ is an eigenvector of T with eigenvalue $\lambda \in F$, then v is also an eigenvector of T^* with eigenvalue $\overline{\lambda}$.

(b) Eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

Proof. Note that $(T - \lambda I)^* = T^* - \overline{\lambda} I$, and that T is normal if and only if $T - \lambda I$ is normal. Suppose $T(v) = \lambda v$. Since T is normal, we have

$$0 = ||(T - \lambda I)(v)|| = \langle v, (T^* - \overline{\lambda} I)(T - \lambda I)(v)\rangle = ||(T^* - \overline{\lambda} I)(v)||.$$

Part (a) follows.

For part (b), suppose λ and μ are distinct eigenvalues with associated eigenvectors u and v, respectively. So $T(u) = \lambda u$ and $T(v) = \mu v$, and from part (a), $T^*(v) = \overline{\mu} v$. Hence

$$\begin{aligned}(\lambda - \mu)\langle u, v\rangle &= \langle \lambda u, v\rangle - \langle u, \overline{\mu} v\rangle \\ &= \langle T(u), v\rangle - \langle u, T^*(v)\rangle \\ &= 0.\end{aligned}$$

Because $\lambda \neq \mu$, the above equation implies that $\langle u, v\rangle = 0$. □

The next theorem is one of the truly important results from the theory of complex inner product spaces. Be sure to compare it with the Real Spectral Theorem.

Theorem 9.2.3. Complex Spectral Theorem *Let V be a finite dimensional complex inner product space and $T \in \mathcal{L}(V)$. Then T is normal if and only if V has an orthonormal basis consisting of eigenvectors of T.*

Proof. First suppose that V has an orthonormal basis $\mathcal{B}$ consisting of eigenvectors of T, so that $[T]_{\mathcal{B}} = A$ is a diagonal matrix. Then A^* is also diagonal and is the matrix $A^* = [T^*]_{\mathcal{B}}$. Since any two diagonal matrices commute, $AA^* = A^*A$. This implies that T is normal.

For the converse, suppose that T is normal. Since V is a complex vector space, we know (by Schur's Theorem) that there is an orthonormal basis $\mathcal{B} = (e_1, \ldots, e_n)$ for which $A = [T]_{\mathcal{B}}$ is upper triangular. If $A = (a_{ij})$, then $a_{ij} = 0$ whenever $i > j$. Also $T(e_1) = a_{11}e_1$, so

$$||T(e_1)||^2 = |a_{11}|^2$$

$$||T^*(e_1)||^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2.$$

Because T is normal, $||T(e_1)|| = ||T^*(e_1)||$. So the two equations above imply that all entries in the first row of A, except possibly the diagonal entry a_{11}, equal 0. It now follows that $T(e_2) = a_{12}e_1 + a_{22}e_2 = a_{22}e_2$, so

$$||T(e_2)||^2 = |a_{22}|^2,$$

$$||T^*(e_2)||^2 = |a_{22}|^2 + |a_{23}|^2 + \cdots + |a_{2n}|^2.$$

Because T is normal, $||T(e_2)|| = ||T^*(e_2)||$. Thus the two equations just above imply that all the entries in the second row of A, except possibly the diagonal entry a_{22}, must equal 0. Continuing in this fashion we see that all the nondiagonal entries of A equal 0, i.e., A is diagonal. □

Corollary 9.2.4. *Let A be a normal, $n \times n$ complex matrix. Then there is a unitary matrix P such that*

$$P^{-1}AP \text{ is diagonal.}$$

Theorem 9.2.5. *Let A be $n \times n$. Then A is normal if and only if the eigenspaces of AA^* are A-invariant.*

Proof. First suppose that A is normal, so $AA^* = A^*A$, and suppose that $AA^*\vec{x} = \lambda\vec{x}$. We show that $A\vec{x}$ also belongs to λ for AA^* : $AA^*(A\vec{x}) = A(AA^*\vec{x}) = A \cdot \lambda\vec{x} = \lambda(A\vec{x})$. For the converse, suppose the eigenspaces of AA^* are A-invariant. We want to show that A is normal. We start with the easy case.

Lemma 9.2.6. *Suppose $BB^* = diag(\lambda_1, \ldots, \lambda_k, 0 \ldots, 0)$ is a diagonal matrix, and suppose that the eigenspaces of BB^* are B-invariant. Then B is normal.*

Proof of Lemma: $\vec{u} = (0, \ldots, 0, u_{k+1}, \ldots, u_n)^T$ is a typical element of the null space of BB^*,i.e., the eigenspace belonging to the eigenvalue 0. First note that the bottom $n-k$ rows of B must be zero, since the (i,i) entry of BB^* is the inner product of the ith row of B with itself, which must be 0 if $i \geq k+1$. So by hypothesis $B(0, \ldots, 0, u_{k+1}, \ldots, u_n)^T = (0, \ldots, 0, v_{k+1}, \ldots, v_n)^T$. Since the top k entries of $B(0, \ldots, 0, u_{k+1}, \ldots, u_n)^T$ must be zero, the entries in the top k rows and last $n-k$ columns must be zero, so

$$B = \begin{pmatrix} B_1 & 0 \\ 0 & 0 \end{pmatrix},$$

where B_1 is $k \times k$ with rank k.

For $1 \leq i \leq k$, the standard basis vector $\vec{e}_i$ is an eigenvector of BB^* belonging to λ_i. And by hypothesis, $BB^* \cdot B\vec{e}_i = \lambda_i B\vec{e}_i$. So $B \cdot BB^*\vec{e}_i = B \cdot \lambda_i\vec{e}_i$, implying that $[B^2B^* - BB^*B]\vec{e}_i = \vec{0}$. But this latter equality is also seen to hold for $k+1 \leq i \leq n$. Hence $B^2B^* = BB^*B$. Now using block multiplication, $B_1^2B_1^* = B_1B_1^*B_1$, implying that $B_1B_1^* = B_1^*B_1$. But this implies that $BB^* = B^*B$. So B is normal, proving the lemma.

Now return to the general case. Let $\vec{u}_1, \ldots, \vec{u}_n$ be an orthonormal basis of eigenvectors of AA^*. Use these vectors as the columns of a matrix U, so that $U^*(AA^*)U = diag(\lambda_1, \ldots, \lambda_k, 0 = \lambda_{k+1}, \ldots, 0 = \lambda_n)$, where $0 \neq \lambda_1 \cdots \lambda_k$ and $k = rank(A)$. Our hypothesis says that if $AA^*\vec{u}_j = \lambda_j A\vec{u}_j$, then $AA^* \cdot A\vec{u}_j = \lambda_j\vec{u}_j$. Put $B = U^*AU$. So $BB^* = U^*AUU^*A^*U = U^*(AA^*)U =$

$diag(\lambda_1, \ldots, \lambda_n)$. Compute $BB^*(U^*\vec{u}_j) = U^*AA^*\vec{u}_j = U^* \cdot \lambda_j\vec{u}_j = \lambda_j U^*\vec{u}_j$. So $U^*\vec{u}_j = \vec{e}_j$ is an eigenvector of BB^* belonging to λ_j.

Also $(BB^*)BU^*\vec{u}_j = U^*AUU^*A^*UU^*AUU^*\vec{u}_j = U^*AA^*A\vec{u}_j = U^* \cdot \lambda_j A\vec{u}_j = \lambda_j \cdot U^*AUU^*\vec{u}_j = \lambda_j B(U^*\vec{u}_j)$. So the eigenspaces of BB^* are B-invariant. Since BB^* is diagonal, by the Lemma we know B is normal. But now it follows easily that $A = UBU^*$ must be normal □

9.3 Decomposition of Real Normal Operators

Throughout this section V is a real inner product space.

Lemma 9.3.1. *Suppose $dim(V) = 2$ and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) T is normal but not self-adjoint.

(b) The matrix of T with respect to every orthonormal basis of V has the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, with $b \neq 0$.

(c) The matrix of T with respect to some orthonormal basis of V has the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, with $b > 0$.

Proof. First suppose (a) holds and let $\mathcal{S} = (e_1, e_2)$ be an orthonormal basis of V. Suppose

$$[T]_{\mathcal{S}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Then $||T(e_1)||^2 = a^2 + b^2$ and $||T^*||^2 = a^2 + c^2$. Because T is normal, by Theorem 9.2.1 $||T(e_1)|| = ||T^*(e_1)||$. Hence $b^2 = c^2$. Since T is not self-adjoint, $b \neq c$, so we have $c = -b$. Then $[T^*]_{\mathcal{S}} = \begin{pmatrix} a & b \\ -b & d \end{pmatrix}$. So $TT^* = \begin{pmatrix} a^2+b^2 & ab-bd \\ ab-bd & b^2+d^2 \end{pmatrix}$, and $T^*T = \begin{pmatrix} a^2+b^2 & -ab+bd \\ -ab+bd & b^2+d^2 \end{pmatrix}$. Since T is normal it follows that $b(a-d) = 0$. Since T is not self-adjoint, $b \neq 0$, implying that $a = d$, completing the proof that (a) implies (b).

Now suppose that (b) holds and let $\mathcal{B} = (e_1, e_2)$ be any orthonormal basis of V. Then either $\mathcal{B}$ or $\mathcal{B}' = (e_1, -e_2)$ will be a basis of the type needed to show that (c) is satisfied.

Finally, suppose that (c) holds, i.e., there is an orthonormal basis $\mathcal{B} = (e_1, e_2)$ such that $[T]_{\mathcal{B}}$ has the form given in (c). Clearly $T \neq T^*$, but a

simple computation with the matrices representing T and T^* shows that $TT^* = T^*T$, i.e., T is normal, implying that (a) holds. □

Theorem 9.3.2. *Suppose that $T \in \mathcal{L}(V)$ is normal and U is a T-invariant subspace. Then*

(a) $U^\perp$ is T-invariant.
(b) U is T^*-invariant.
(c) $(T|_U)^* = (T^*)|_U$.
(d) $T|_U$ is a normal operator on U.
(e) $T|_{U^\perp}$ is a normal operator on $U^\perp$.

Proof. Let $\mathcal{B}' = (e_1, \ldots, e_m)$ be an orthonormal basis of U and extend it to an orthonormal basis $\mathcal{B} = (e_1, \ldots, e_m, f_1, \ldots, f_n)$ of V. Since U is T-invariant,

$$[T]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}, \text{ where } A = [T|_U]_{\mathcal{B}'}.$$

For each j, $1 \leq j \leq m$, $||T(e_j)||^2$ equals the sum of the squares of the absolute values of the entries in the jth column of A. Hence

$$\sum_{j=1}^{m} ||T(e_j)||^2 = \begin{array}{c}\text{the sum of the squares of the absolute} \\ \text{values of the entries of } A.\end{array} \tag{9.3}$$

For each j, $1 \leq j \leq m$, $||T^*(e_j)||^2$ equals the sum of the squares of the absolute values of the entries in the jth rows of A and B. Hence

$$\sum_{j=1}^{m} ||T^*(e_j)||^2 = \begin{array}{c}\text{the sum of the squares of the absolute} \\ \text{values of the entries of } A \text{ and } B.\end{array} \tag{9.4}$$

Because T is normal, $||T(e_j)|| = ||T^*(e_j)||$ for each j. It follows that the entries of B must all be 0, so

$$[T]_{\mathcal{B}} = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}.$$

This shows that $U^\perp$ is T-invariant, proving (a).

But now we see that

$$[T^*]_{\mathcal{B}} = \begin{pmatrix} A^* & 0 \\ 0 & C^* \end{pmatrix},$$

implying that U is T^*-invariant. This completes a proof of (b).

Now let $S = T|_U$. Fix $v \in U$. Then

$$\langle S(u), v\rangle = \langle T(u), v\rangle = \langle u, T^*(v)\rangle \ \forall u \in U.$$

Because $T^*(v) \in U$ (by (b)), the equation above shows that $S^*(v) = T^*(v)$, i.e., $(T|_U)^* = (T^*)|_U$, completing the proof of (c). Parts (d) and (e) now follow easily. □

At this point the reader should review the concept of block multiplication for matrices partitioned into blocks of the appropriate sizes. In particular, if A and B are two *block diagonal* matrices each with k blocks down the diagonal, with the jth block being $n_j \times n_j$, then the product AB is also block diagonal. Suppose that A_j, B_j are the jth blocks of A and B, respectively. Then the jth block of AB is A_jB_j:

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & A_m \end{pmatrix} \cdot \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & B_m \end{pmatrix} = \begin{pmatrix} A_1B_1 & 0 & \cdots & 0 \\ 0 & A_2B_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & A_mB_m \end{pmatrix}.$$

We have seen the example $T(x, y) = (-y, x)$ of an operator on $\mathcal{R}^2$ that is normal but has no eigenvalues, so has no diagonal matrix. However, the following theorem says that normal operators have block-diagonal matrices with blocks of size at most 2 by 2.

Theorem 9.3.3. *Suppose that V is a real inner product space and $T \in \mathcal{L}(V)$. Then T is normal if and only if there is an orthonormal basis of V with respect to which T has a block diagonal matrix where each block is a 1-by-1 matrix or a 2-by-2 matrix of the form*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \tag{9.5}$$

with $b > 0$.

Proof. First suppose that V has an orthonormal basis $\mathcal{B}$ for which $[T]_{\mathcal{B}}$ is block diagonal of the type described in the theorem. Since a matrix of the form given in Eq. 9.5 commutes with its adjoint, clearly T is also normal.

For the converse, suppose that T is normal. Our proof proceeds by induction on the dimension n of V. For $n = 1$ the result is obvious. For $n = 2$ if T is self-adjoint it follows from the Real Spectral Theorem; if T is not self-adjoint, use Lemma 9.3.1. Now assume that $n = \dim(V) > 2$ and that the desired result holds on vector spaces of dimension smaller than n. By Theorem 7.3.1 we may let U be a T-invariant subspace of dimention 1 if there is one. If there is not, then we let U be a 2-dimensional T-invariant subspace. First, if $\dim(U) = 1$, let e_1 be a nonzero vector in U with norm 1. So $\mathcal{B}' = (e_1)$ is an orthonormal basis of U. Clearly the matrix $[T|_U]_{\mathcal{B}'}$ is 1-by-1. If $\dim(U) = 2$, then $T|_U$ is normal (by Theorem 9.3.2, but not self-adjoint (since otherwise $T|_U$, and hence T, would have an eigenvector in U by Lemma 9.1.4). So we may choose an orthonormal basis of U with respect to which the matrix of $T|_U$ has the desired form. We know that $U^\perp$ is T-invariant and $T|_{U^\perp}$ is a normal operator on $U^\perp$. By our induction hypothesis there is an orthonormal basis of $U^\perp$ of the desired type. Putting together the bases of U and $U^\perp$ we obtain an orthonormal basis of V of the desired type. □

9.4 Positive Operators

In this section V is an inner product space. An operator $T \in \mathcal{L}(V)$ is said to be a *positive operator* provided

$$T = T^* \text{ and } \langle T(v), v\rangle \geq 0 \ \forall v \in V.$$

Note that if V is a complex space, then having $\langle T(v), v\rangle \in \mathcal{R}$ for all $v \in V$ is sufficient to force T to be self-adjoint (by Theorem 9.9.1, part (c)). So $\langle T(v), v\rangle \geq 0$ for all $v \in V$ is sufficient to force T to be positive.

There are many examples of positive operators. If P is any orthogonal projection, then P is positive. (You should verify this!) In the proof of Lemma 9.1.3 we showed that if $T \in \mathcal{L}(V)$ and if $\alpha, \beta \in \mathcal{R}$ are such that $\alpha^2 < 4\beta$, then $T^2+\alpha T+\beta I$ is positive. You should think about the analogy between positive operators (among all operators) and the nonnegative real numbers among all complex numbers. This will be made easier by the theorem that follows, which collects the main facts about positive operators.

If $S, T \in \mathcal{L}(V)$ and $S^2 = T$, we say that S is a *square root* of T.

Theorem 9.4.1. *Let $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) T is positive;
(b) T is self-adjoint and all the eigenvalues of T are nonnegative:
(c) T has a positive square root;
(d) T has a self-adjoint square root;
(e) There exists an operator $S \in \mathcal{L}(V)$ such that $T = S^*S$.

Proof. We will prove the (a) $\implies$ (b) $\implies$ (c) $\implies$ (d) $\implies$ (e) $\implies$ (a). Suppose that (a) holds, i.e., T is positive, so in particular T is self-adjoint. Let v be a nonzero eigenvector belonging to the eigenvalue λ. Then

$$0 \leq \langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda \langle v, v\rangle,$$

implying that λ is a nonnegative number, so (b) holds.

Now suppose that (b) holds, so T is self-adjoint and all the eigenvalues of T are nonnegative. By the Real and Complex Spectral Theorems, there is an orthonormal basis $\mathcal{B} = (e_1, \ldots, e_n)$ of V consisting of eigenvectors of T. Say $T(e_i) = \lambda_j e_i$, where each $\lambda_i \geq 0$. Define $S \in \mathcal{L}(B)$ by

$$S(e_j) = \sqrt{\lambda_j} e_j, \quad 1 \leq j \leq n.$$

Since S has a real diagonal matrix with respect to the basis $\mathcal{B}$, clearly S is self-adjoint. Now suppose $v = \sum_{j=1}^n a_j e_j$. Then $\langle S(v), v\rangle = \sum_{j=1}^n \sqrt{\lambda_j}|a_j|^2 \geq 0$, so S is a positive square root of T. This shows (b) $\implies$ (c).

Clearly (c) implies (d), since by definition every positive operator is self-adjoint. So suppose (d) holds and let S be a self-adjoint operator with $T = S^2$. Since S is self-adjoint, $T = S^*S$, proving that (e) holds.

Finally, suppose that $T = S^*S$ for some $S \in \mathcal{L}(V)$. It is easy to check that $T^* = T$, and then $\langle T(v), v\rangle = \langle S^*S(v), v\rangle = \langle (S(v), S(v)\rangle \geq 0$, showing that (e) implies (a). $\square$

An operator can have many square roots. For example, in addition to $\pm I$ being square roots of I, for each $a \in F$ and for each nonzero $b \in F$, we have that

$$A = \begin{pmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{pmatrix} \text{ satisfies } A^2 = I.$$

However, things are different if we restrict our attention to positive operators.

Theorem 9.4.2. *Each positive operator on V has a unique positive square root.*

Proof. Let $T \in \mathcal{L}(V)$ be positive with nonnegative distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Since T is self-adjoint, we know by Theorem 7.2.8 (and the Spectral Theorems) that

$$V = \text{null}(T - \lambda_1 I) \oplus \cdots \oplus \text{null}(T - \lambda_m I).$$

By the preceding theorem we know that T has a positive square root S. Suppose α is an eigenvalue of S. If $v \in \text{ null}(S - \alpha I)$. Then $T(v) = S^2(v) = \alpha^2 v$, so α^2 is some eigenvalue of T, i.e., $\alpha = \sqrt{\lambda_i}$ for some i. Clearly

$$\text{null}(S - \sqrt{\lambda_j} \subseteq \text{ null}(T - \lambda_j I).$$

Since the only possible eigenvalues of S are $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m}$, and because S is self-adjoint, we also know that

$$V = \text{ null}(S - \sqrt{\lambda_1} I) \oplus \cdots \oplus \text{ null}(S - \sqrt{\lambda_m} I).$$

A dimension argument then shows that

$$\text{null}(S - \sqrt{\lambda_j} I) = \text{ null}(T - \lambda_j I)$$

for each j. In other words, on $\text{null}(T - \lambda_j I)$, the operator S is just multiplication by $\sqrt{\lambda_j}$. Thus S, the positive square root of T, is uniquely determined by T. □

9.5 Isometries

An operator $S \in \mathcal{L}(V)$ is called an *isometry* provided

$$||S(v)|| = ||v|| \ \forall v \in V.$$

For example, λI is an isometry whenever $\lambda \in F$ satisfies $|\lambda| = 1$. More generally, suppose that $\lambda_1, \ldots, \lambda_n$ are scalars with absolute value 1 and $S \in \mathcal{L}(V)$ satisfies $S(e_j) = \lambda_j e_j$ for some orthonormal basis $\mathcal{B} = (e_1, \ldots, e_n)$ of V. For $v \in V$ we have

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n \tag{9.6}$$

and (using the Pythagorean theorem)

$$||v||^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2. \tag{9.7}$$

Apply S to both sides of Eq. s9.6:

$$S(v) = \lambda_1\langle v, e_1\rangle e_1 + \cdots + \lambda_n\langle v, e_n\rangle e_n.$$

This last equation along with $|\lambda_j| = 1$ shows that

$$||S(v)||^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2. \tag{9.8}$$

If we compare Eqs. 9.7 and 9.8, we see that $||v|| = ||S(v)||$, i.e., S is an isometry. In fact, this is the prototypical isometry, as we shall see. The next theorem collects the main results concerning isometries.

Theorem 9.5.1. *Suppose V is an n-dimensional inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) S is an isometry;

(b) $\langle S(u), S(v)\rangle = \langle u, v\rangle \ \forall u, v \in V$;

(c) $S^*S = I$;

(d) $(S(e_1), \ldots, S(e_n))$ is an orthonormal basis of V whenever $(e_1, \ldots, e_n)$ is an orthonormal basis of V;

(e) there exists an orthonormal basis $(e_1, \ldots, e_n)$ of V for which $(S(e_1), \ldots, S(e_n))$ is orthonormal;

(f) S^* is an isometry;

(g) $\langle S^*(u), S^*(v)\rangle = \langle u, v\rangle \ \forall u, v \in V$;

(h) $SS^* = I$.

(i) $\langle S^*(e_1), \ldots, S^*(e_n)\rangle$ is orthonormal whenever $(e_1, \ldots, e_n)$ is an orthonormal list of vectors in V;

(j) there exists an orthonormal basis $(e_1, \ldots, e_n)$ of V for which $(S^*(e_1), \ldots, S^*(e_n))$ is orthonormal.

Proof. To start, suppose S is an isometry. If V is a real inner-product space, then for all $u, v \in V$, using the real polarization identity we have

$$\begin{aligned}
\langle S(u), S(v)\rangle &= \frac{||S(u) + S(v)||^2 - ||S(u) - S(v)||^2}{4} \\
&= \frac{||S(u + v)||^2 - ||S(u - v)||^2}{4} \\
&= \frac{||u + v||^2 - ||u - v||^2}{4} \\
&= \langle u, v\rangle.
\end{aligned}$$

If V is a complex inner product space, use the complex polarization identity in the same fashion. In either case we see that (a) implies (b).

Now suppose that (b) holds. Then

$$\begin{aligned}\langle (S^*S - I)(u), v\rangle &= \langle S(u), S(v)\rangle - \langle u, v\rangle \\ &= 0\end{aligned}$$

for every $u, v \in V$. In particular, if $v = (S^*S - I)(u)$, then necessarily $(S^*S - I)(u) = 0$ for all $u \in V$, forcing $S^*S = I$. Hence (b) implies (c).

Suppose that (c) holds and let $(e_1, \ldots, e_n)$ be an orthonormal list of vectors in V. Then

$$\langle S(e_j), S(e_k)\rangle = \langle S^*S(e_j), e_k\rangle = \langle e_j, e_k\rangle.$$

Hence $(S(e_1), \ldots, S(e_n))$ is orthonormal, proving that (c) implies (d).

Clearly (d) implies (e).

Suppose that $(e_1, \ldots, e_n)$ is a basis of V for which $(S(e_1), \ldots, S(e_n))$ is orthonormal. For $v \in V$,

$$\begin{aligned}||S(v)||^2 &= ||S(\sum_{i=1}^{n}\langle v, e_i\rangle e_i||^2 \\ &= ||\sum_{i=1}^{n}||\langle v, e_i\rangle S(e_i)||^2 \\ &= \sum_{i=1}^{n}|\langle v, e_i\rangle|^2 \\ &= ||v||^2.\end{aligned}$$

Taking square roots we see that (e) implies (a).

We have now shown that (a) through (e) are equivalent. Hence replacing S by S^* we have that (f) through (j) are equivalent. Clearly (c) and (h) are equivalent, so the proof of the theorem is complete. □

The preceding theorem shows that an isometry is necessarily normal (see parts (a), (c) and (h)). Using the characterization of normal operators proved earlier we can now give a complete description of all isometries. But as usual, there are separate statements for the real and the complex cases.

Theorem 9.5.2. *Let V be a (finite dimensional) complex inner product space and let $S \in \mathcal{L}(V)$. Then S is an isometry if and only if there is an orthonormal basis of V consisting of eigenvectors of S all of whose corresponding eigenvalues have absolute value 1.*

Proof. The example given at the beginning of this section shows that the condition given in the theorem is sufficient for S to be an isometry. For the converse, suppose that $S \in \mathcal{L}(V)$ is an isometry. By the complex spectral theorem there is an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors of S. For $1 \leq j \leq n$, let λ_j be the eigenvalue corresponding to e_j. Then

$$|\lambda_j| = ||\lambda_j e_j|| = ||S(e_j)|| = ||e_j|| = 1.$$

Hence each eigenvalue of S has absolute value 1, completing the proof. □

The next result states that every isometry on a real inner product space is the direct sum of pieces that look like rotations on 2-dimensional subpaces, pieces that equal the identity operator, and pieces that equal multiplication by -1. It follows that an isometry on an odd-dimensional real inner product space must have 1 or -1 as an eigenvalue.

Theorem 9.5.3. *Suppose that V is a real inner product space and $S \in \mathcal{L}(V)$. Then S is an isometry if and only if there is an orthonormal basis of V with respect to which S has a block diagonal matrix where each block on the diagonal is a 1-by-1 matrix containing 1 or -1, or a 2-by-2 matrix of the form*

$$\begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}, \quad \textit{with } \theta \in (0, \pi). \tag{9.9}$$

Proof. First suppose that S is an isometry. Because S is normal, there is an orthonormal basis of V such that with respect to this basis S has a block diagonal matrix, where each block is a 1-by-1 matrix or a 2-by-2 matrix of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{with } b > 0. \tag{9.10}$$

If λ is an entry in a 1-by-1 block along the diagonal of the matrix of S (with respect to the basis just mentioned), then there is a basis vector e_j such that $S(e_j) = \lambda e_j$. Because S is an isometry, this implies that $|\lambda| = 1$ with λ real, forcing $\lambda = \pm 1$.

Now consider a 2-by-2 matrix of the form in Eq. 9.10 along the diagonal of the matrix of S. There are basis vectors $e_j,\ e_{j+1}$ such that

$$S(e_j) = ae_j + be_{j+1}.$$

Thus

$$1 = ||e_j||^2 = ||S(e_j)||^2 = a^2 + b^2.$$

This equation, along with the condition that $b > 0$, implies that there exists a number $\theta \in (0, \pi)$ such that $a = cos(\theta)$ and $b = sin(\theta)$. Thus the matrix in Eq. 9.10 has the required form. This completes the proof in one direction.

Conversely, suppose that there is an orthonormal basis of V with respect to which the matrix of S has the form stated in the theorem. There there is a direct sum decomposition

$$V = U_1 \oplus \cdots \oplus U_m,$$

where each subspace U_j is a subspace of V having dimension 1 or 2. Furthermore, any two vectors belonging to distinct U's are orthogonal, and each $S|_{U_j}$ is an isometry mapping U_j into U_j. If $v \in V$, write

$$v = \sum_{i=1}^{m} u_i,\ u_j \in U_j.$$

Applying S to this equation and taking norms gives

$$\begin{aligned} ||S(v)||^2 &= ||S(u_1) + \cdots + S(u_m)||^2 \\ &= ||S(u_1)||^2 + \cdots + ||S(u_m)||^2 \\ &= ||u_1||^2 + \cdots + ||u_m||^2 \\ &= ||v||^2. \end{aligned}$$

This shows that S is an isometry, as desired. □

9.6 The Polar Decomposition

Theorem 9.6.1. Polar Decomposition *If $T \in \mathcal{L}(V)$, then there exists an isometry $S \in \mathcal{L}(V)$ such that*

$$T = S \circ \sqrt{T^*T}.$$

Proof. Let $T \in \mathcal{L}(V)$. Then for each $v \in V$, $||T(v)||^2 = \langle T(v), T(v)\rangle = \langle T^*T(v), v\rangle = \langle \sqrt{T^*T}(v), \sqrt{T^*T}(v)\rangle = ||\sqrt{T^*T}(v)||^2$.

So we have established

$$||T(v)|| = ||\sqrt{T^*T}(v)|| \ \forall v \in V. \tag{9.11}$$

The next step is to construct a map

$$S_1 : \mathrm{Im}(\sqrt{T^*T}) \to \mathrm{Im}(T) : \sqrt{T^*T}(v) \mapsto T(v),$$

and show that it is a well-defined isometry.

$$||T(v_1) = T(v_2)|| = ||T(v_1 - v_2)|| = ||\sqrt{T^*T}(v_1 - v_2)|| = ||\sqrt{T^*T}(v_1) - \sqrt{T^*T}(v_2)||.$$

This shows that $T(v_1) = T(v_2)$ if and only if $\sqrt{T^*T}(v_1) = \sqrt{T^*T}(v_2)$. In fact it shows that S_1 is well-defined and is a bijection from $Im(\sqrt{T^*T})$ onto $Im(T)$. One consequence of this is that

$$\begin{aligned} \dim(\mathrm{Im}(\sqrt{T^*T})) &= \dim(\mathrm{Im}(T)) \text{ and} \\ \dim(\mathrm{Im}(\sqrt{T^*T}))^\perp &= \dim(\mathrm{Im}(T))^\perp. \end{aligned}$$

It is also easy to check that S_1 is linear. Moreover, if $v = \sqrt{T^*T}(u)$, then $||S_1(v)|| = ||T(u)|| = ||\sqrt{T^*T}(u)|| = ||v||$, implying that S_1 is an isometry.

Now construct an orthonormal basis $(e_1, \ldots, e_m)$ of $(\mathrm{range}(\sqrt{T^*T})^\perp$ and an orthonormal basis $(f_1, \ldots, f_m)$ of $(\mathrm{Im}(T))^\perp$. Define

$$S_2 : (\mathrm{Im}(\sqrt{T^*T})^\perp \to (\mathrm{Im}(T))^\perp$$

by: $S_2(e_j) = f_j$ (and extend linearly). It follows that $||S_2(w)|| = ||w||$ for all $w \in (\mathrm{Im}(\sqrt{T^*T})^\perp$. Here S_2 is an isometry by part (e) of Theorem 9.5.1.

We know that

$$V = \mathrm{Im}(\sqrt{T^*T}) \perp (\mathrm{Im}(\sqrt{T^*T}))^\perp.$$

For $v \in V$, write $v = u + w$ with $u \in \mathrm{Im}(\sqrt{T^*T})$, $w \in (\mathrm{Im}(\sqrt{T^*T}))^\perp$. Define $S : V \to V$ by $S(v) = S_1(u) + S_2(w)$. It is easy to check that $S \in \mathcal{L}(V)$. Moreover, $||S(v)||^2 = ||S_1(u) + S_2(w)||^2 = ||S_1(u)||^2 + ||S_2(w)||^2 = ||u||^2 + ||w||^2 = ||v||^2$, implying S is an isometry. (Here we used the fact that $S_1(u) \in \mathrm{Im}(T)$ and $S_2(w) \in (\mathrm{Im}(T))^\perp$.) The only thing left to check is that $T = S \circ \sqrt{T^*T}$, but this is obvious by the way S_1 is defined on the image of $\sqrt{T^*T}$. $\square$

We can visualize this proof as follows. Start with $T \in \mathcal{L}(V)$.

$$\begin{array}{ccccc} \mathrm{Im}(\sqrt{T^*T}) & \perp & \left(\mathrm{Im}(\sqrt{T^*T})\right)^{\perp} & & e_1, \ldots, e_n \\ \sqrt{T^*T}(v) & & & & e_j \\ S_1 \downarrow & & S_2 \downarrow & & \downarrow \\ T(v) & & & & f_j \\ \mathrm{Im}(T) & \perp & (\mathrm{Im}(T))^{\perp} & & f_1, \ldots, f_n \end{array}$$

$$S = S_1 \oplus S_2 \implies S \text{ is an isometry.}$$

And

Polar Decomposition: $T = S \circ \sqrt{T^*T}$.

If T is invertible, then $S = T \circ (\sqrt{t^*T})^{-1}$ is unique. If T is not invertible, then $\mathrm{Im}(T) \neq V$, so $S_2 \neq -S_2$. Hence $S' = S_1 \oplus -S_2$ yields s a polar decomposition of T distinct from that given by S'.

The polar decomposition states that each operator on V can be written as the product of an isometry and a positive operator. Thus we can write each operator on V as the product of two operators, each of which is of a type that we have completely describe and understand reasonably well. We know there is an orthonormal basis of V with respect to which the isometry S has a diagonal matrix (if $F = \mathcal{C}$) or a block diagonal matrix with blocks of size at most 2-by-2 (if $F = \mathcal{R}$), and there is an orthonormal basis of V with respect to which $\sqrt{T^*T}$ has a diagonal matrix. Unfortunately, there may not be one orthonormal basis that does both at the same time. However, we can still say something interesting. This is given by the singular value decomposition which is discussed in the next section.

9.7 The Singular-Value Decompostition

Statement of the General Result

Let F be a subfield of the complex number field $\mathcal{C}$, and let $A \in M_{m,n}(F)$. Clearly A^*A is self-adjoint, so there is an orthonormal basis $(v_1, \ldots, v_n)$ for F^n (whose elements we view as column vectors) consisting of eigenvectors of A^*A with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $\langle\ ,\ \rangle$ denote the usual

inner product on F^n. Since $||Av_i||^2 = \langle Av_i, Av_i\rangle = \langle A^*Av_i, v_i\rangle = \langle \lambda_i v_i, v_i\rangle = \lambda_i ||v_i||^2 = \lambda_i$, we have proved the following lemma.

Lemma 9.7.1. *With the notation as above,*

(i) Each eigenvalue λ_i of A^*A is real and nonnegative. Hence WLOG we may assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$.

(ii) With $s_i = \sqrt{\lambda_i}$ for $1 \leq i \leq n$, we say that the s_i, $1 \leq i \leq n$, are the singular values *of A. Hence the singular values of A are the lengths of the vectors $Av_1, \ldots, Av_n$.*

This is enough for us to state the theorem concerning the Singular Value Decomposition of A.

Theorem 9.7.2. *Given $A \in M_{m,n}(F)$ as above, there are unitary matrices $U \in M_m(F)$ and $V \in M_n(F)$ such that*

$$A = U\Sigma V^*, \text{ where } \Sigma = \begin{pmatrix} s_1 & & & 0 \\ & \ddots & & \\ & & s_r & \\ & & & 0 \\ & & & 0 \end{pmatrix} \tag{9.12}$$

is a diagonal $m \times n$ matrix and $s_1 \geq s_2 \geq \cdots \geq s_r$ are the positive (i.e., nonzero) singular values of A.

(i) The columns of U are eigenvectors of AA^*, the first r columns of U form an orthonormal basis for the column space $col(A)$ of A, and the last $m - r$ columns of U form an orthonormal basis for the (right) null space of the matrix A^*.

(ii) The columns of V are eigenvectors of A^*A, the last $n - r$ columns of V form an orthonormal basis for the (right) null space $null(A)$ of A, and the first r columns of V form an orthonormal basis for the column space $col(A^*)$.

(iii) The rank of A is r.

Proof. Start by supposing that $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$. Since $||Av_i|| = \sqrt{\lambda_i} = s_i$, clearly $Av_i \neq 0$ iff $1 \leq i \leq r$. Also, if $i \neq j$, then $\langle Av_i, Av_j\rangle = \langle A^*Av_i, v_j\rangle = \lambda_i\langle v_i, v_j\rangle = 0$. So $(Av_1, \ldots, Av_n)$ is an orthogonal list. Moreover, $Av_i \neq \overline{0}$ iff $1 \leq i \leq r$, since $||Av_i|| = \sqrt{\lambda_i} = s_i$. So clearly $(Av_1, \ldots, Av_r)$ is a linearly independent list that spans a subspace of $col(A)$. On the other hand suppose that $y = Ax \in col(A)$. Then $x = \sum c_i v_i$, so $y = Ax = \sum_{i=1}^n c_i Av_i = \sum_{i=1}^r c_i Av_i$. It follows that $(Av_1, \ldots Av_r)$ is a basis

for $col(A)$, implying that r =dim$(col(A))$ = rank(A). Of course, then the right null space of A has dimension $n - r$.

For $1 \leq i \leq r$, put $u_i = \frac{Av_i}{||Av_i||} = \frac{1}{s_i} \cdot Av_i$, so that $Av_i = s_i \cdot u_i$. Now extend $(u_1, \ldots, u_r)$ to an orthonormal basis $(u_1, \ldots, u_n)$ of F^m. Let $U = [u_1, \ldots, u_m]$ be the matrix whose jth column is u_j. Similarly, put $V = [v_1, \ldots, v_n]$, so U and V are unitary matrices. And

$$AV = [Av_1, \ldots, Av_r, \overline{0}, \ldots, \overline{0}] = [s_1 u_1, \ldots, s_r u_r, \overline{0}, \ldots, \overline{0}] =$$

$$= [u_1, \ldots, u_m] \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ & & s_r \\ 0 & & \ddots \end{pmatrix} = U\Sigma,$$

where Σ is diagonal, $m \times n$, with the singular values of A along the diagonal.

Then $AV = U\Sigma$ implies $A = U\Sigma V^*$, $A^* = V\Sigma^* U^*$, so $A^*A = V\Sigma^* U^* \cdot U\Sigma V^* = V\Sigma^*\Sigma V^*$.

Then

$$(A^*A)V = V\Sigma^*\Sigma V^* V = V\Sigma^*\Sigma = V \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix} = (\lambda_1 v_1, \ldots, \lambda_r v_r, \overline{0}, \cdots, \overline{0}).$$

Since $(A^*A)v_j = \overline{0}$ for $j > r$, implying $v_j^* A^* A v_j = 0$, forcing $Av_j = \overline{0}$, it is clear that $v_{r+1}, \ldots, v_n$ form an orthonormal basis for $null(A)$. Since $\lambda_j v_j = (A^*/a)v_j = A^* \sqrt{\lambda_j} u_j$, we see $A^* u_j = s_j v_j$. It now follows readily that $v_1, \ldots, v_r$ form a basis for the column space of A^*.

Similarly, $AA^* = U\Sigma\Sigma^* U^*$, so that

$$(AA^*)U = U\Sigma\Sigma^* = (\lambda_1 u_1, \ldots, \lambda_r u_r, \overline{0}, \cdots, \overline{0}).$$

It follows that $(u_1, \ldots, u_n)$ consists of eigenvectors of AA^*. Also, $u_{r+1}, \ldots, u_m$ form an orthonormal basis for the (right) null space of A^*, and $u_1, \ldots, u_r$ form an orthonormal basis for $col(A)$. □

Recapitulation

We want to practice recognizing and/or finding a singular value decomposition for a given $m \times n$ matrix A over the subfield F of $\mathcal{C}$.

$$A = U\Sigma V^* \text{ if and only if } A^* = V\Sigma^* U^*. \tag{9.13}$$

Here Σ is $m \times n$ and diagonal with the nonzero singular values $s_1 \geq s_2 \geq \ldots, s_r$ down the main diagonal as its only nonzero entries. Similarly, Σ^* is $n \times m$ and diagonal with $s_1 \geq \cdots \geq s_r$ down the main diagonal as its only nonzero entries. So the nonzero eigenvalues of A^*A are identical to the nonzero eigenvalues of AA^*. The only difference is the multiplicity of 0 as an eigenvalue.

$$\text{For } 1 \leq i \leq r,\ Av_i = s_i u_i \text{ and } A^* u_i = s_i v_i. \tag{9.14}$$

$$(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n) \text{ is an orthonormal basis of eigenvectors of } A^*A. \tag{9.15}$$

$$(u_1, \ldots, u_r, u_{r+1}, \ldots, u_m) \text{ is an orthonormal basis of eigenvectors of } AA^*. \tag{9.16}$$

$$(v_1, \ldots, v_r) \text{ is a basis for } col(A^*) \text{ and } (v_{r+1}, \ldots, v_n) \text{ is a basis for } null(A). \tag{9.17}$$

$$(u_1, \ldots, u_r) \text{ is a basis for } col(A) \text{ and } (u_{r+1}, \ldots, u_m) \text{ is a basis for } null(A^*). \tag{9.18}$$

We can first find $(v_1, \ldots, v_n)$ as an orthonormal basis of eigenvectors of A^*A, put $u_i = \frac{1}{s_i} \cdot Av_i$, for $1 \leq i \leq r$, and then complete $(u_1, \ldots, u_r)$ to an orthonormal basis $(u_1, \ldots, u_m)$ of eigenvectors of AA^*. Sometimes here it is efficient to use the fact that $(u_{r+1}, \ldots, u_m)$ form a basis for the null space of AA^* or of A^*. If $n < m$, this is the approach usually taken.

Alternatively, we can find $(u_1, \ldots, u_m)$ as an orthonormal basis of eigenvectors of AA^* (with eigenvalues ordered from largest to smallest), put $v_i = \frac{1}{s_i} A^* u_i$ for $1 \leq i \leq r$, and then complete $(v_1, \ldots, v_r)$ to an orthonormal basis $(v_1, \ldots, v_n)$ of eigenvectors of A^*A. If $m < n$, this is the approach usually taken.

9.7.3 Two Examples

Problem 1. Find the Singular Value Decomposition of $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$.

Solution: $A^*A = A^TA = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$. In this case it is easy to find an orthonormal basis of R^2 consisting of eigenvectors of A^*A. Put $v_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $V = [v_1, v_2]$.

Then $A^*A[v_1, v_2] = [18v_1, 0 \cdot v_2]$. So put $u_1 = \frac{Av_1}{||Av_1||} = \frac{1}{3\sqrt{2}} \begin{pmatrix} -\sqrt{2} \\ 2\sqrt{2} \\ -2\sqrt{2} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}$ = the first column of U.

It is now easy to see that $w_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $w_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ form a basis of ${u_1}^{\perp}$. We apply Gram-Schmidt to (w_1, w_2) to obtain

$$u_2 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} \frac{-2}{\sqrt{45}} \\ \frac{4}{\sqrt{5}} \\ \frac{5}{\sqrt{45}} \end{pmatrix}.$$

Now put $U = [u_1, u_2, u_3]$.

It follows that $A = U\Sigma V^*$ is one singular value decomposition of A, where $\Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Problem 2. Compute a singular value decomposition of $A = \begin{pmatrix} 1 & 0 & i \\ 0 & 1 & -i \end{pmatrix}$.

Solution: In this case AA^* is 2×2, while A^*A is 3×3, so we start with $AA^* = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. Here AA^* has eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 1$. So the singular values of A^* (and hence of A) are $s_1 = \sqrt{3}$ and $s_2 = 1$. It follows that $\Sigma = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Also $\Sigma^* = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.

In this case we choose to compute an orthonormal basis (u_1, u_2) of eigenvectors of AA^*. It is simple to check that $AA^* - 3I = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$, and we may take $u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}$. Similarly, $AA^* - I = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, and we may take $u_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. Then $U = [u_1, u_2]$. At this point we must put

$$v_1 = \frac{1}{s_1} A^* u_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{-2i}{\sqrt{6}} \end{pmatrix},$$

and

$$v_2 = 1 \cdot A^* u_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

At this point we know that (v_3) must be an orthonormal basis for $\{\text{span}(v_1, v_2)\}^\perp$. So we want $v_3 = (x, y, z)^T$ with $0 = \langle (1, -1, -2i), (x, y, z) \rangle = \overline{x} - \overline{y} - 2i\overline{z}$, and $0 = \langle (1, 1, 0), (x, y, z) \rangle = \overline{x} + \overline{y}$. It follows easily that $(\overline{x}, \overline{y}, \overline{z}) = (i\overline{z}, -i\overline{z}, \overline{z})$, where we must choose z so that the norm of this vector is 1. If we put $\overline{z} = \frac{-i}{\sqrt{3}}$, i.e., $z = \frac{i}{\sqrt{3}}$, then $(x, y, z) = (\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{i}{\sqrt{3}})$.

Then

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2i}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{i}{\sqrt{3}} \end{pmatrix},$$

which is a singular value decomposition of A.

THE REMAINDER OF THIS SECTION MAY BE CONSIDERED TO BE STARRED.

Definition If λ is an eigenvalue of the matrix A, then $\dim(\text{null}(T-\lambda I))$ is called the *geometric multiplicity of the eigenvalue* λ.

Theorem 9.7.4. *Let M be $m_2 \times m_1$ and N be $m_1 \times m_2$. Put*

$$A = \begin{pmatrix} 0 & N \\ M & 0 \end{pmatrix}.$$

Then the following are equivalent:

(i) $\lambda \neq 0$ is an eigenvalue of A with (geometric) multiplicity f.

(ii) $-\lambda \neq 0$ is an eigenvalue of A with (geometric) multiplicity f.

(iii) $\lambda^2 \neq 0$ is an eigenvalue of MN with (geometric) multiplicity f.

(iv) $\lambda^2 \neq 0$ is an eigenvalue of NM with (geometric) multiplicity f.

Proof. Step 1. Show (i) $\leftrightarrow$ (ii) Let $AU = \lambda U$ for some matrix U of rank f. Write $U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$, and put $\tilde{U} = \begin{pmatrix} U_1 \\ -U_2 \end{pmatrix}$, where U_i has m_i rows for $i = 1, 2$. Then $AU = \lambda U$ becomes

$$\begin{pmatrix} 0 & N \\ M & 0 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} NU_2 \\ MU_1 \end{pmatrix} = \lambda \begin{pmatrix} U_1 \\ U_2 \end{pmatrix},$$

so $NU_2 = \lambda U_1$ and $MU_1 = \lambda U_2$.
This implies $A\tilde{U} = \begin{pmatrix} 0 & N \\ M & 0 \end{pmatrix} \begin{pmatrix} U_1 \\ -U_2 \end{pmatrix} = \begin{pmatrix} -\lambda U_1 \\ \lambda U_2 \end{pmatrix} = -\lambda \tilde{U}$. Since $rank(U) = rank(\tilde{U})$, the first equivalence follows.

Step 2. Show (iii) $\leftrightarrow$ (iv). Let $MNU' = \lambda^2 U'$ for some matrix U' of rank f. Then $(NM)(NU') = \lambda^2 NU'$, and $rank(NU') = rank(U')$, since $rank(\lambda^2 U') = rank(MNU') \leq rank(U')$, and $\lambda \neq 0$. So NM has λ^2 as eigenvalue with geometric multiplicity at least f. Interchanging roles of N and M proves (iii) $\leftrightarrow$ (iv).

Step 3. Show (i) $\leftrightarrow$ (iii) Let $AU = \lambda U$ with U having rank f, and $\tilde{U} = \begin{pmatrix} U_1 \\ -U_2 \end{pmatrix}$ as in Step 1. Put $n = m_1 + m_2$. Then $A^2\Big(U; \tilde{U}\Big) = \lambda^2\Big(U; \tilde{U}\Big)$. since $NU_2 = \lambda U_1$ and $MU_1 = \lambda U_2$, U_1 and U_2 have the same row space. So $row(U) = row(U_1) = row(U_2)$. This implies $rank(U_1) = rank(U_2) = f$.

Using column operations we can transform $\begin{pmatrix} U_1 & U_1 \\ U_2 & -U_2 \end{pmatrix}$ into $\begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}$, which has rank $2f$. So λ^2 is an eigenvalue of A^2 with geometric multiplicity at least $2f$.

On the other hand, the geometric multiplicity of an eigenvalue $\lambda^2 \neq 0$ of A^2 equals $n - rank(A^2 - \lambda^2 I) = n - rank((A - \lambda I)(A + \lambda I)) \leq n + n - rank(A - \lambda I) - rank(A + \lambda I) = 2f$. □

Note: The **singular values** of a complex matrix N are sometimes defined to be the positive eigenvalues of $\begin{pmatrix} 0 & N \\ N^* & 0 \end{pmatrix}$. By the above result we see that they are the same as the positive square roots of the nonzero eigenvalues of NN^* (or of N^*N) as we have defined them above.

9.8 Pseudoinverses and Least Squares*

Let A be an $m \times n$ matrix over F (where F is either $\mathcal{R}$ or $\mathcal{C}$) and suppose the rank of A is r. Let

$$A = U\Sigma V^* \text{ be a singular value decomposition of } A.$$

So Σ is an $m \times n$ diagonal matrix with the nonzero singular values $s_1 \geq s_2 \geq \ldots, \geq s_r$ down the main diagonal as its only nonzero entries. Define Σ^+ to be the $n \times m$ diagonal matrix with diagonal equal to $(\frac{1}{s_1}, \frac{1}{s_2}, \ldots, \frac{1}{s_r}, 0, \ldots)$. Then both $\Sigma\Sigma^+$ and $\Sigma^+\Sigma$ have the general block form

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

but the first product is $m \times m$ and the second is $n \times n$.

Definition The *Moore-Penrose generalized inverse of* A (sometimes just called the *pseudoinverse* of A) is the $n \times m$ matrix A^+ over F defined by

$$A^+ = V\Sigma^+U^*.$$

Theorem 9.8.1. *Let A be an $m \times n$ matrix over F with pseudoinverse A^+ (as defined above). Then the following three properties hold:*

(a) $AA^+A = A$;

(b) $A^+AA^+ = A^+$;

(c) AA^+ and A^+A are hermitian (i.e., self-adjoint).

Proof. All three properties are easily shown to hold using the definition of A^+. You should do this now. □

Our definition of "the" Moore-Penrose generalized inverse would not be valid if it were possible for there to be more than one. However, the following theorem shows that there is at most one (hence exactly one!) such pseudoinverse.

Theorem 9.8.2. *Given an $m \times n$ matrix A, there is at most one matrix satisfying the three properties of A^+ given in Theorem 9.8.1. This means that A has a unique pseudoinverse.*

Proof. Let B and C be pseudoinverses of A. i.e., satisfying the three properties of A^+ in Theorem 9.8.1. Then

$$CA = C(ABA) = CA(BA)^* = CAA^*B^* = (A(CA)^*)^*B^* = (ACA)^*B^*$$

$$= A^*B^* = (BA)^* = BA,$$

i.e.,

$$CA = BA.$$

Then

$$B = BAB = B(AB)^* = BB^*A^*,$$

so that

$$B = BB^*(ACA)^* = BB^*A^*(AC)^* = BAC = CAC = C.$$

□

At this point we want to review and slightly revise the Gram-Schmidt algorithm given earlier. Let $(v_1, \ldots, v_n)$ be any list of column vectors in F^m. Let these be the columns of an $m \times n$ matrix A. Let $W = span(v_1, \ldots, v_n)$ be the column space of A. Define $u_1 = v_1$. Then for $2 \leq i \leq n$ put

$$u_i = v_i - \alpha_{1i}u_1 - \alpha_{2i}u_2 - \cdots - \alpha_{i-1,i}u_{i-1},$$

where $\alpha_{ji} = \frac{\langle v_i, u_j \rangle}{\langle u_j, u_j \rangle}$, if $u_j \neq \vec{0}$, and $\alpha_{ji} = 0$ if $u_j = \vec{0}$. Hence

$$v_i = \alpha_{1i}u_1 + \alpha_{2i}u_2 + \alpha_{3i}u_3 + \cdots + \alpha_{i-1,i}u_{i-1} + u_i. \tag{9.19}$$

Let Q_0 be the $m \times n$ matrix whose columns are $u_1, u_2, \ldots, u_n$, respectively, and let R_0 be the $n \times n$ matrix given by

$$R_0 = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & 1 & \cdots & \alpha_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{9.20}$$

The u_i in Q_0 are constructed by the Gram-Schmidt method and form an orthogonal set. This means that Q_0 has orthogonal columns, some of which may be zero. Let Q and R be the matrices obtained by deleting the zero columns from Q_0 and the corresponding rows from R_0, and by dividing each nonzero column of Q_0 by its norm and multiplying each corresponding row of R_0 by that same norm. Then Eq. 9.20 becomes

$$A = QR \text{ with } R \text{ upper triangular and } Q \text{ having orthonormal columns.} \tag{9.21}$$

Eq. 9.21 is the *normalized QR-decomposition* of A. Note that if A has rank k, then Q is $m \times k$ with rank k and R is $k \times n$ and upper triangular with rank k. The columns of Q form an orthonormal basis for the column space W of A.

If we compute Q^*Q, since the columns of Q are orthonormal, we get

$$Q^*Q = I_k.$$

Since $(u_1, \ldots, u_k)$ is an orthonormal basis for W, if $P_0 \in \mathcal{L}(F^n)$ is the orthogonal projection onto W, then for $v \in F^n$ we have

$$P_0(v) = \sum_{i=1}^{k} \langle v, u_i \rangle u_i = (u_1, \ldots, u_k) \begin{pmatrix} \langle v, u_1 \rangle \\ \vdots \\ \langle v, u_k \rangle \end{pmatrix} = Q \cdot \begin{pmatrix} u_1^* \\ \vdots \\ u_k^* \end{pmatrix} \cdot v = QQ^*v.$$

So QQ^* is the *projection matrix* projecting v onto $W = \mathrm{col}(Q)$. Hence QQ^*v is the unique vector in W closest to v.

Lemma 9.8.3. *Suppose that the $m \times n$ matrix A has rank k and that $A = BC$, where B is $m \times k$ with rank k and C is $k \times n$ with rank k. Then*

$$A^{++} = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$$

is the pseudoinverse of A.

Proof. It is rather straightforward to verify that the three properties of Theorem 9.8.1 are satisfied by this A^{++}. Then by Theorem 9.8.2 we know that a matrix A^+ satisfying the three properties of Theorem 9.8.1 is uniquely determined by these properties.) □

Corollary 9.8.4. *If $A = QR$ is a normalized QR-decomposition of A, then $A^{+++} = R^*(RR^*)^{-1}Q^*$ is the pseudoinverse of A.*

Given that $A = U\Sigma V^*$ is a singular value decomposition of A, partition the two matrices as follows: $U = (U_k, U_{m-k})$ and $V = (V_k, V_{n-k})$, where U_k is $m \times k$ and its columns are the first k columns of U, i.e., they form an orthonormal basis of col(A). Similarly, V_k is $k \times n$ and it columns are the first k columns of V, so they form an orthonormal basis for col(A^*). This is the same as saying that the rows of V_k^* form an orthonormal basis for row(A). Now let D be the $k \times k$ diagonal matrix with the nonzero singular values of A down the diagonal, i.e., $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$. Now using block multiplication we see that

$$A = (U_k, U_{m-k}) \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_k^* \\ V_{n-k}^* \end{pmatrix} = U_k D_k V_k^*.$$

This expression $A = U_k D_k V_k^*$ is called the *reduced singular value decomposition* of A. Here U_k is $m \times k$ with rank k, and $D_k V_k^*$ is $k \times n$ with rank k. Then $A = QR = U_k \cdot D_k V_k^*$ where $Q = U_k$ is $m \times k$ with rank k, and $R = D_k V_k^*$ is $k \times n$ with rank k. The columns of Q form an orthonormal basis for col(A), and the rows of R form an orthogonal basis for row(A). So a pseudoinverse $A^+$is given by

$$A^+ = R^*(RR^*)^{-1}(Q^*Q)^{-1}Q^* = V_k D_k^{-1} U_k^*, \text{ after some computation.}$$

Suppose we are given the equation

$$Ax = b, \tag{9.22}$$

where A is $m \times n$ and b is $m \times 1$. It is possible that this equation is not consistent. In this case we want to find $\hat{x}$ so that $A\hat{x}$ is as close to b as possible, i.e., it should be the case that $A\hat{x}$ is the projection $\hat{b}$ of b onto the column space of A. Put

$$\hat{x} = A^+ b = V_k D_k^{-1} U_k^* b.$$

Then

$$\begin{aligned} A\hat{x} &= (U_k D_k V_k^*) V_k D_k^{-1} U_k^* b \\ &= U_k D D^{-1} U_k^* b \\ &= U_k U_k^* b, \end{aligned} \tag{9.23}$$

which by the paragraph preceding Lemma 9.8.3 must be the projection $\hat{b}$ of b onto $\mathrm{col}(A)$. Thus $\hat{x}$ is a "least-squares solution" to $Ax = b$, in the sense that $\hat{x}$ minimizes $||A\hat{x} - b||$ (which is the square root of a sum of squares). Moreover, it is true that $\hat{x}$ has the smallest length among all least-squares solutions to $Ax = b$.

Theorem 9.8.5. *Let A be $m \times n$ with rank k. Let $A = U\Sigma V^*$ be a normalized singular value decomposition of A. Let $U = (U_k, U_{m-k})$ and $V = (V_k, V_{n-k})$ be partitioned as above. Then $A^+ = V_k D_k^{-1} U_k^*$, as above, and $\hat{x} = A^+ b$ is a least-squares solution to $Ax = b$. If x_0 is any other least-squares solution, i.e., $||Ax_0 - b|| = \|A\hat{x} - b||$, then $||\hat{x}|| \leq ||x_0||$, with equality if and only if $x_0 = \hat{x}$.*

Proof. What remains to be shown is that if $||Ax_0 - b|| = \|A\hat{x} - b||$, then $||\hat{x}|| \leq ||x_0||$, with equality if and only if $x_0 = \hat{x}$. We know that $||Ax - b||$ is minimized precisely when $Ax = \hat{b}$ is the projection $U_k U_k^* b$ of b onto $\mathrm{col}(A)$. So suppose $Ax_0 = \hat{b} = A\hat{x}$, which implies $A(x_0 - \hat{x}) = 0$, i.e., $x_0 - \hat{x} \in \mathrm{null}(A)$. By Eq. 9.17 this means that $x_0 = \hat{x} + V_{n-k} z$ for some $z \in F^{n-k}$. We claim that $\hat{x} = A^+ b = V_k D_k^{-1} U_k^* b$ and $V_{n-k} z$ are orthogonal. For, $\langle V_k D_K^{-1} U_k^* b, V_{n-k} z \rangle = z^*(V_{n-k}^* V_k) D_k^{-1} U_K^* b = 0$, because by Eq. 9.15, $V_{n-k}^* V_k = 0_{(n-k)\times k}$. Then $||x_0||^2 = ||\hat{x} + V_{n-k} z||^2 = ||\hat{x}||^2 + ||V_{n-k} z||^2 \geq ||\hat{x}||^2$, with equality if and only if $||V_{n-k} z||^2 = 0$ which is if and only if $V_{n-k} z = \vec{0}$ which is if and only if $x_0 = \hat{x}$. □

Theorem 9.8.6. *Let A be $m \times n$ with rank k. Then A has a unique (Moore-Penrose) pseudoinverse. If $k = n$, $A^+ = (A^*A)^{-1} A^*$, and A^+ is a left inverse of A. If $k = m$, $A^+ = A^*(AA^*)^{-1}$, and A^+ is a right inverse of A. If $k = m = n$, $A^+ = A^{-1}$.*

Proof. We have already seen that A has a unique pseudoinverse. If $k = n$, $A = B$, $C = I_k$ yields a factorization of the type used in Lemma 9.8.3, so $A^+ = (A^*A)^{-1} A^*$. Similarly, if $k = m$, put $B = I_k$, $C = A$. Then $A^+ = A^*(AA^*)^{-1}$. And if A is invertible, so that $(A^*A)^{-1} = A^{-1}(A^*)^{-1}$, by the $k = n$ case we have $A^+ = (A^*A)^{-1} A^* = A^{-1}(A^*)^{-1} A^* = A^{-1}$. □

9.9 Norms, Distance and More on Least Squares*

In this section the norm $\|A\|$ of an $m \times n$ matrix A over $\mathcal{C}$ is defined by:

$$\|A\| = (tr(A^*A))^{1/2} = \left(\sum_{i,j} |A_{ij}^2 \right)^{1/2}.$$

(Here we write $tr(A)$ for the trace of A.)

Note: This norm is the 2-norm of A thought of as a vector in the mn-dimensional vector space $M_{m,n}(\mathcal{C})$. It is almost obvious that $\|A\| = \|A^*\|$.

Theorem 9.9.1. *Let A, P, Q be arbitrary complex matrices for which the appropriate multiplications are defined. Then*

$$\|AP\|^2 + \|(I - AA^+)Q\|^2 = \|AP + (I - AA^+)Q\|^2.$$

Proof.

$$\|AP + (I - AA^+)Q\|^2 = tr\left\{[AP + (I - AA^+)Q]^*[AP + (I - AA^+)Q]\right\}$$

$$= \|AP\|^2 + tr\left\{(AP)^*(I - AA^+)Q\right\} + tr\left\{((I - AA^+)Q)^*AP\right\} + \|(I - AA^+)Q\|^2.$$

We now show that each of the two middle terms is the trace of the zero matrix. Since the second of these matrices is the conjugate transpose of the other, it is necessary only to see that one of them is the zero matrix. So for the first matrix:

$$(AP)^*(I - AA^+)Q = (AP)^*(I - AA^+)^*Q = ((I - AA^+)AP)^*Q =$$

$$= (AP - AA^+AP)^*Q = (AP - AP)^*Q = 0.$$

□

Theorem 9.9.2. *Let A and B be $m \times n$ and $m \times p$ complex matrices. Then the matrix $X_0 = A^+B$ enjoys the following properties:*

(i) $\|AX - B\| \geq \|AX_0 - B\|$ for all $n \times p$ S. Moreover, equality holds if and only if $AX = AA^+B$.

(ii) $\|X\| \geq \|X_0\|$ for all X such that $\|AX - B\| = \|AX_0 - B\|$. with equality if and only if $X = X_0$.

Proof.

$$\|AX - B\|^2 =$$

$$= \|A(X - A^+B) + ((I - AA^+)(-B)\|^2 = \|A(X - A^+B\|^2 + \|(I - AA^+)(-B)\|^2$$

$$= \|AX - AA^+B\|^2 + \|AA^+B - B\|^2 \geq \|AA^+B - B\|^2,$$

with equality if and only if $AX = AA^+B$. This completes the proof of (i). Now interchange A^+ and A in the preceding theorem and assume $AX = AA^+B$ so that equality holds in (i). Then we have

$$\|X\|^2 = \|A^+B + X - A^+(AA^+B)\|^2 = \|A^+B + (I - A^+A)X\|^2$$

$$= \|A^+B\|^2 + \|X - A^+AX\|^2 = \|A^+B\|^2 + \|X - A^+B\|^2 \geq \|A^+B\|^2,$$

with equality if and only if $X = A^+B$. □

NOTE: Suppose we have n (column) vectors $v_1, \ldots, v_n$ in $\mathcal{R}^m$ and some vector y in $\mathcal{R}^m$, and we want to find that vector in the space $\text{span}(v_1, \ldots, v_n) = V$ which is closest to the vector y. First, we need to use only an independent subset of the v_i's. So use row-reduction techniques to find a basis $(w_1, \ldots, w_k)$ of V. Now use these w_i to form the columns of a matrix A. So A is $m \times k$ with rank k, and the pseudoinverse of A is simpler to compute than if we had used all of the v_i's to form the columns of A. Put $x_0 = A^+y$. Then $Ax_0 = AA^+y$ is the desired vector in V closest to y. Now suppose we want to find the distance from y to V, i.e., the distance

$$d(y, AA^+y) = \|y - AA^+y\|.$$

It is easier to compute the square of the distance first:

$$\|y - AA^+y\|^2 = \|(I - AA^+)y\|^2 = y^*(I - AA^+)^*(I - AA^+)y =$$

$$= y^*(I - AA^+)(I - AA^+)y = y^*(I - AA^+ - AA^+AA^+AA^+)y = y^*(I - AA^+)y.$$

We have proved the following:

Theorem 9.9.3. *The distance between a vector y of $\mathcal{R}^m$ and the column space of an $m \times n$ matrix A is $(y^*(I - AA^+)y)^{\frac{1}{2}}$.*

Theorem 9.9.4. *$AXB = C$ has a solution X if and only if $AA^+CB^+B = C$, in which case for any Y,*

$$X = A^+CB^+ + Y - A^+AYBB^+ \text{ is a solution.}$$

This gives the general solution.

Proof. If $AXB = C$, then $C = AXB = AA^+AXBB^+B = AA^+CB^+B$. Conversely, if $C = AA^+CB^+B$, then $X = A^+CB^+$ is a particular solution of $AXB = C$. Any expression of the form $X = Y - A^+AYBB^+$ satisfies $AXB = 0$. And if $AXB = 0$, then $X = Y - A^+AYBB^+$ for $Y = X$. Put $X_0 = A^+CB^+$ and let X_1 be any other solution to $AXB = C$. Then $X = X_1 - X_0$ satisfies $AXB = 0$, so that

$$X_1 - X_0 = X = Y - A^+AYBB^+ \text{ for some } Y.$$

□

If $x = (x_1, \ldots, x_n)^T$ and $y = (y_1, \ldots, y_n)^T$ are two points of $\mathcal{C}^n$, the *distance* between x and y is defined to be

$$d(x, y) = \|x - y\| = \left(\sum |x_i - y_i|^2\right)^{1/2}.$$

A *hyperplane* H in $\mathcal{C}^n$ is the set of vectors $x = (x_1, \ldots, x_n)^T$ satisfying an equation of the form $\sum_{i=1}^n a_i x_i + d = 0$, i.e.,

$$H = \{x \in \mathcal{C}^n : Ax + d = 0\}, \text{ where } A = (a_1, \ldots, a_n) \neq \vec{0}.$$

$A = 1 \cdot (a_1, \ldots, a_n)$, where 1 is 1×1 of rank 1 and $(a_n, \ldots, a_n)$ is $1 \times n$ of rank 1, so

$$A^+ = A^*(AA^*)^{-1} = \frac{1}{\|A\|^2}A^*.$$

For such an A you should now show that $\|A^+\| = \frac{1}{\|A\|}$.

Let $y = (y_1, \ldots, y_n)^T$ be a given point of $\mathcal{C}^n$ and let $H : Ax + d = 0$ be a given hyperplane. We propose to find the point x_0 of H that is closest to y_0 and to find the distance $d(x_0, y_0)$.

Note that x is on H if and only if $Ax + d = 0$ if and only if $A(x - y_0) - (-d - Ay_0) = 0$. Hence our problem is to find x_0 such that

(i) $A(x_0 - y_0) - (-d - Ay_0) = 0$, (i.e., x_0 is on H),

and

(ii) $\|x_0 - y_0\|$ is minimal (so x_0 is the point of H closest to y_0).

Theorem 9.9.2 says that the vector $x = x_0 - y_0$ that satisfies $Ax - (-d - Ay_0) = 0$ with $\|x\|$ minimal is given by

$$x = x_0 - y_0 = A^+(-d - Ay_0) = \frac{1}{\|A\|^2}A^*(-d - Ay_0).$$

Hence

$$x_0 = y_0 + \frac{-d}{\|A\|^2}A^* - \frac{1}{\|A\|^2}A^*Ay_0.$$

And

$$d(y_0, H) = d(y_0, x_0) = \|x_0 - y_0\| = \|A^+(-d - Ay_0)\|$$

$$= |d + Ay_0| \cdot \|A^+\| = \frac{|Ay_0 + d|}{\|A\|}.$$

Note: This distance formula generalizes well known formulas for the distance from a point to a line in $\mathcal{R}^2$ and from a point to a plane in $\mathcal{R}^3$.

We now recall the method of approximating by least squares. Suppose y is a function of n real variables $t^{(1)}, \ldots, t^{(n)}$, and we want to approximate y as a linear function of these variables. this means we want to find (real) numbers $x_0, \ldots, x_n$ for which

1. $x_0 + x_1t^{(1)} + \cdots + x_nt^{(n)}$ is as "close" to the function $y = y(t^{(1)}, \ldots, t^{(n)})$ as possible.

   Suppose we have m measurements of y corresponding to m different sets of values of the $t^{(j)}$: $(y_i; t_i^{(1)}, \ldots, t_i^{(n)})$, $i = 1, \ldots, m$. The problem is to determine $x_0, \ldots, x_n$ so that

2. $y_i = x_1t_i^{(1)} + \cdots + x_nt^{(n)} + r_i$, $1 \leq i \leq m$, where the r_i's are small in some sense.

   Put $t_i^{(j)} = a_{ij}$, $y = (y_1, \ldots, y_m)^T$, $x = (x_0, \ldots, x_n)^T$, $r = (r_1, \ldots, r_m)^T$, and

$$A = \begin{pmatrix} 1 & a_{11} & \cdots & a_{1n} \\ 1 & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & a_{m_1} & \cdots & a_{mn} \end{pmatrix}.$$

Then 2. becomes:

3. $y = Ax + r$.

   A standard interpretation of saying that "r is small" is to say that for some some weighting constants $w_1, \ldots, w_m$, the number $S = \sum_{i=1}^m w_i r_i^2 = r^T W r$, where $W = \mathrm{diag}(w_1, \ldots, w_m)$ is minimal.

   As $S = \sum_{i=1}^m w_i(y_i - s_0 - x_1 a_{i1} - x_2 a_{i2} - \cdots - x_n a_{in})^2$, to minimize S as a function of $x_0, \ldots, x_n$, we require that $\frac{\partial S}{\partial x_k} = 0$ for all k.

   Then

$$\frac{\partial S}{\partial x_0} = -2 \sum_i w_i(y_i - x_0 - x_1 a_{i1} - \cdots - x_n a_{in}) = 0$$

   implies

4. $$(\sum_{i=1}^m w_i)x_0 + (\sum_{i=1}^m w_i a_{i1})x_1 + \cdots + (\sum_{i=1}^m w_i a_{in})x_n = \sum w_i y_i.$$

   And for $1 \le k \le n$: $\frac{\partial S}{\partial x_k} = -2 \sum_{i=1}^m w_i(y_i - x_0 - x_1 a_{i1} - \cdots x_n a_{in}) a_{ik} = 0$ implies

5. $$(\sum_{i=1}^m w_i a_{ik})x_0 + (\sum_{i=1}^m w_i a_{i1} a_{ik})x_1 + \cdots + (\sum_{i=1}^m w_i a_{in} a_{ik})x_n = \sum w_i y_i a_{ik}.$$

   It is easy to check that

$$A^T W = \begin{pmatrix} 1 & \cdots & 1 \\ a_{11} & \cdots & a_{m1} \\ \vdots & \cdots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} w_1 & \cdots 0 & \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_m \end{pmatrix}$$

$$= \begin{pmatrix} w_1 & \cdots & w_m \\ a_{11}w_1 & \cdots & a_{m1}w_m \\ \vdots & \cdots & \vdots \\ a_{1n}w_n & \cdots & a_{mn}w_m \end{pmatrix}.$$

So putting together the right hand sides of 4. and 5., we obtain

6.
$$(A^TW)y = \begin{pmatrix} \sum w_iy_i \\ \sum a_{i1}w_iy_i \\ \vdots \\ \sum a_{in}w_iy_i \end{pmatrix}.$$

Also we compute

$$A^TWA = \begin{pmatrix} w_1 & \cdots & w_m \\ a_{11}w_1 & \cdots & a_{m1}w_m \\ \vdots & \vdots & \vdots \\ a_{1n}w_1 & \cdots & a_{mn}w_m \end{pmatrix} \begin{pmatrix} 1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Comparing the above with the left hand sides of 4. and 5., and putting the above equations together, we obtain the following system of $n+1$ equations in $n+1$ unknowns:

7. $(A^TWA)x = A^TWy$.

We now reconsider this problem taking advantage of our results on generalized inverses. So A is a real $m \times n$ matrix, y is a real $m \times 1$ matrix, and $x \in R^n$ is sought for which $(y-Ax)^T(y-Ax) = \|Ax-y\|^2$ is minimal (here we put $W = I$). Theorem 9.9.2 solves this problem by putting $x = A^+y$. We show that this also satisfies the above condition $A^TAx = A^Ty$, as should be expected. For suppose $A = BC$ with B $m \times k$, C $k \times n$, where $k = \text{rank}(A) = \text{rank}(B) = \text{rank}(C)$. Then

$$A^TAy = A^TAA^+y = C^TB^TBCC^T(CC^T)^{-1}(B^TB)^{-1}B^Ty = C^TB^Ty = A^Ty.$$

So when $W = I$ the generalized inverse solution does the following: First, it actually constructs a solution of the original problem that, second, has the smallest norm of any solution.

9.10 Exercises

1. Prove that if $T \in \mathcal{L}(V)$ is normal, then $\text{Im}(T) = \text{Im}(T^*)$.

2. Prove that if $T \in \mathcal{L}(V)$ is normal, then

$$\text{null}(T^k) = \text{null}(T) \quad \text{and} \quad \text{Im}(T^k) = \text{Im}(T)$$

   for every positive integer k.

3. Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.

4. Prove that $A^+ = A^{-1}$ when A is nonsingular.

5. Determine all singular value decompositions of I with $U = V$. Show that all of them lead to exactly the same pseudoinverse.

6. Prove:

   (a) The rank of A^+ is the same as the rank of A.

   (b) If A is self-conjugate, then A^+ is self-conjugate.

   (c) $(cA)^+ = \frac{1}{c}A^+$ for $c \neq 0$.

   (d) $(A^+)^* = (A^*)^+$.

   (e) $(A^+)^+ = A$.

   (f) Show by a counterexample that in general $(AB)^+ \neq B^+A^+$.

   (g) If A is $m \times r$, B is $r \times n$, and both matrices have rank k, then $(AB)^+ = B^+A^+$.

7. Suppose that x is $m \times 1$ and y is $1 \times m$. Compute

   (a) x^+; (b) y^+; (c) $(xy)^+$.

   CHECK OUT PAGE 159 IN AXLER.

Chapter 10

Decomposition WRT a Linear Operator

To start this chapter we assume that F is an arbitrary field and let V be an n-dimensional vector space over F. Note that this necessarily means that we are not assuming that V is an inner product space.

10.1 Powers of Operators

First note that if $T \in \mathcal{L}(V)$, if k is a nonnegative integer, and if $T^k(v) = \vec{0}$, then $T^{k+1}(v) = \vec{0}$. Thus $\text{null}(T^k) \subseteq \text{null}(T^{k+1})$. It follows that

$$\{\vec{0}\} = \text{null}(T^0) \subseteq \text{null}(T^1) \subseteq \cdots \text{null}(T^k) \subseteq \text{null}(T^{k+1}) \subseteq \cdots . \quad (10.1)$$

Theorem 10.1.1. *Let $T \in \mathcal{L}(V)$ and suppose that m is a nonnegative integer such that $\text{null}(T^m) = \text{null}(T^{m+1})$. Then $\text{null}(T^m) = \text{null}(T^k)$ for all $k \geq m$.*

Proof. Let k be a positive integer. We want to prove that

$$\text{null}(T^{m+k}) = \text{null}(T^{m+k+1}).$$

We already know that $\text{null}(T^{m+k}) \subseteq \text{null}(T^{m+k+1})$. To prove the inclusion in the other direction, suppose that $v \in \text{null}(T^{m+k+1})$. Then $\vec{0} = T^{m+1}(T^k)(v)$. So $T^k(v) \in \text{null}(T^{m+1}) = \text{null}(T^m)$, implying that $\vec{0} = T^m(T^k(v)) = T^{m+k}(v)$. Hence $\text{null}(T^{m+k+1}) \subseteq \text{null}(T^{m+k})$, completing the proof. □

Corollary 10.1.2. *If $T \in \mathcal{L}(V)$, $n = dim(V)$, and k is any positive integer, then $null(T^n) = null(T^{n+k})$.*

Proof. If the set containments in Eq. 10.1 are strict for as long as possible, at each stage the dimension of a given null space is at least one more than the dimension of the preceding null space. Since the entire space has dimension n, at most $n+1$ proper containments can occur. Hence by Theorem 10.1.1 the Corollary must be true. □

Definition: Let $T \in \mathcal{L}(V)$ and suppose that λ is an eigenvalue of T. A vector $v \in V$ is called a *generalized eigenvector* of T corresponding to λ provided

$$(T - \lambda I)^j(v) = \vec{0} \text{ for some positive integer } j. \tag{10.2}$$

By taking $j = 1$ we see that each eigenvector is also a generalized eigenvector. Also, the set of generalized eigenvectors is a subspace of V. Moreover, by Corollary 10.1.2 (with T replaced by $T - \lambda I$) we see that

Corollary 10.1.3. *The set of generalized eigenvectors corresponding to λ is exactly equal to $null[(T - \lambda I)^n]$, where $n = dim(V)$.*

An operator $N \in \mathcal{L}(V)$ is said to be *nilpotent* provided some power of N is the zero operator. This is equivalent to saying that the minimal polynomial $p(x)$ for N has the form $p(x) = x^j$ for some positive integer j, implying that the characteristic polynomial for N is x^n. Then by the Cayley-Hamilton Theorem we see that N^n is the zero operator.

Now we turn to dealing with images of operators. Let $T \in \mathcal{L}(V)$ and $k \geq 0$. If $w \in \text{Im}(T^{k+1})$, say $w = T^{k+1}(v)$ for some $v \in V$, then $w = T^k(T(v) \in \text{Im}(T^k)$. In other words we have

$$V = \text{Im}(T^0) \supseteq \text{Im}T^1 \supseteq \cdots \supseteq \text{Im}(T^k) \supseteq \text{Im}(T^{k+1}) \cdots . \tag{10.3}$$

Theorem 10.1.4. *If $T \in \mathcal{L}(V)$, $n = dim(V)$, and k is any positive integer, then*

$$Im(T^n) = Im(T^{n+k}).$$

Proof. We use the corresponding result already proved for null spaces.

$$\begin{aligned} \dim(\text{Im}(T^{n+k})) &= n - \dim(\text{null}(T^{n+k})) \\ &= n - \dim(\text{null}(T^n)) \\ &= \dim(\text{Im}(T^n)). \end{aligned}$$

Now the proof is easily finished using Eq. 10.3 □

10.2 The Algebraic Multiplicity of an Eigenvalue

If the matrix A is upper triangular, so is the matrix $xI - A$, whose determinant is the product of its diagonal elements and also is the characteristic polynomial of A. So the number of times a given scalar λ appears on the diagonal of A is also the algebraic multiplicity of λ as a root of the characteristic polynomial of A. The next theorem states that the algebraic multiplicity of an eigenvalue is also the dimension of the space of generalized eigenvectors associated with that eigenvalue.

Theorem 10.2.1. *Let $n = dim(V)$, let $T \in \mathcal{L}(V)$ and let $\lambda \in F$. Then for each basis $\mathcal{B}$ of V for which $[T]_{\mathcal{B}}$ is upper triangular, λ appears on the diagonal of $[T]_{\mathcal{B}}$ exactly dim(null$[(T - \lambda I)^n]$) times.*

Proof. For notational convenience, we first assume that $\lambda = 0$. Once this case is handled, the general case is obtained by replacing T with $T - \lambda I$. The proof is by induction on n, and the theorem is clearly true when $n = 1$. We assume that $n > 1$ and that the theorem holds on spaces of dimension $n - 1$.

Suppose that $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of V for which $[T]_{\mathcal{B}}$ is the upper triangular matrix

$$\begin{pmatrix} \lambda_1 & & & * \\ & \ddots & & \\ & & \lambda_{n-1} & \\ 0 & & & \lambda_n \end{pmatrix}. \tag{10.4}$$

Let $U = \text{span}(v_1, \ldots, v_n)$. Clearly U is invariant under T and the matrix of $T|_U$ with respect to the basis $(v_1, \ldots, v_{n-1})$ is

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_{n-1} \end{pmatrix}. \tag{10.5}$$

By our induction hypothesis, 0 appears on the diagonal of the matrix in Eq. 10.5 exactly dim(null$((T|_U)^{n-1})$) times. Also, we know that null$((T|_U)^{n-1}) =$ null$((T|_U)^n)$, because dim$(U) = n - 1$. Hence

$$0 \text{ appears on the diagonal of } 10.5 \text{ dim}(\text{null}((T|_U)^n))) \text{ times.} \tag{10.6}$$

The proof now breaks into two cases, depending on whether $\lambda_n = 0$ or not. First consider the case where $\lambda_n \neq 0$. We show in this case that

$$\text{null}(T^n) \subseteq U. \tag{10.7}$$

This will show that $\text{null}(T^n) = \text{null}((T|_U)^n)$, and hence Eq 10.6 will say that 0 appears on the diagonal of Eq. 10.4 exactly $\dim(\text{null}(T^n))$ times, completing the proof in the case where $\lambda_n \neq 0$.

It follows from Eq. 10.4 that

$$[T^n]_{\mathcal{B}} = ([T]_{\mathcal{B}})^n = \begin{pmatrix} \lambda_1^n & & & * \\ & \ddots & & \\ & & \lambda_{n-1}^n & \\ 0 & & & \lambda_n^n \end{pmatrix}. \tag{10.8}$$

This shows that

$$T^n(v_n) = u + \lambda_n^n v_n$$

for some $u \in U$. Suppose that $v \in \text{null}(T^n)$. Then $v = \tilde{u} + av_n$ where $\tilde{u} \in U$ and $a \in F$. Thus

$$\vec{0} = T^n(v) = T^n(\tilde{u}) + aT^n(v_n) = T^n(\tilde{u}) + au + a\lambda_n^n v_n.$$

Because $T^n(\tilde{u})$ and au are in U and $v_n \notin U$, this implies that $a\lambda_n^n = 0$. Since $\lambda_n \neq 0$, clearly $a = 0$. Thus $v = \tilde{u} \in U$, completing the proof of Eq. 10.7, and hence finishing the case with $\lambda_n \neq 0$.

Suppose $\lambda_n = 0$. Here we show that

$$\dim(\text{null}(T^n)) = \dim(\text{null}((T|_U)^n)) + 1, \tag{10.9}$$

which along with Eq. 10.6 will complete the proof when $\lambda_n = 0$.

First consider

$$\begin{aligned} \dim(\text{null}(T^n)) &= \dim(U \cap \text{null}(T^n) + \dim(U + \text{null}(T^n)) - \dim(U) \\ &= \dim(\text{null}((T|_U)^n)) + \dim(U + \text{null}(T^n)) - (n-1). \end{aligned}$$

Also,

$$n = \dim(V) \geq \dim(U + \text{null}(T^n)) \geq \dim(U) = n - 1.$$

It follows that if we can show that $\text{null}(T^n)$ contains a vector not in U, then Eq. 10.9 will be established. First note that since $\lambda_n = 0$, we have $T(v_n) \in U$, hence

$$T^n(v_n) = T^{n-1}(T(v_n)) \in \text{Im}[(T|_U)^{n-1}] = \text{Im}[(T|_U)^n].$$

This says that there is some $u \in U$ for which $T^n(u) = T^n(v_n)$. Then $u - v_n$ is not in U but $T^n(u - v_n) = \vec{0}$. Hence Eq. 10.9 holds, completing the proof. □

At this point we know that the geometric multiplicity of λ is the dimension of the null space of $T - \lambda I$, i.e., the dimension of the eigenspace associated with λ, and this is less than or equal to the algebraic multiplicity of λ, which is the dimension of the null space of $(T-\lambda)^n$ and also equal to the multiplicity of λ as a root of the characteristic polynomial of T, at least in the case that T is upper triangularizable. This is always true if F is algebraically closed. Moreover, in this case the following corollary is clearly true:

Corollary 10.2.2. *If F is algebraically closed, then the sum of the algebraic multiplicities of of all the eigenvalues of T equals $dim(V)$.*

For any $f(x) \in F[x]$, T and $f(T)$ commute, so that the null space of $p(T)$ is invariant under T.

Theorem 10.2.3. *Suppose F is algebraically closed and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, and let $U_1, \ldots, U_m$ be the corresponding subspaces of generalized eigenvectors. Then*

(a) $V = U_1 \oplus \ldots \oplus U_m$;

(b) each U_j is T-invariant;

(c) each $(T - \lambda_j)$ is nilpotent.

Proof. Since $U_j = \text{null}(T - \lambda_j I)^n$ for each j, clearly (b) follows. Clearly (c) follows from the definitions. Also the sum of the multiplicities of the eigenvalues equals $\dim(V)$, i.e.

$$\dim(V) = \sum_{j=1}^{m} \dim(U_j).$$

Put $U = U_1 + \cdots + U_m$. Clearly U is invariant under T. Hence we can define $S \in \mathcal{L}(U)$ by

$$S = T|_U.$$

It is clear that S has the same eigenvalues, with the same multiplicities, as does T, since all the generalized eigenvectors of T are in U. Then the dimension of U is the sum of the dimensions of the generalized eigenspaces of T, forcing $V = U$, and $V = \oplus \sum_{j=1}^{m} U_j$, completing the proof of (a). □

By joining bases of the various generalized eigenspaces we obtain a basis for V conisting of generalized eigenvectors, proving the following corollary.

Corollary 10.2.4. *Let F be algebraically closed and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.*

Lemma 10.2.5. *Let N be a nilpotent operator on a vector space V over any field F. Then there is a basis of V with respect to which the matrix of N has the form*

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}, \tag{10.10}$$

Proof. First choose a basis of null(N). Then extend this to a basis of null(N^2). Then extend this to a basis of null(N^3). Continue in this fashion until eventually a basis of V is obtained, since $V = \text{null}(N^m)$ for sufficiently large m. A little thought should make it clear that with respect to this basis, the matrix of N is upper triangular with zeros on the diagonal. □

10.3 Elementary Operations

For $1 \leq i \leq n$, let e_i denote the column vector with a 1 in the ith position and zeros elsewhere. Then $e_i e_j^T$ is an $n \times n$ matrix with all entries other than the (i, j) entry equal to zero, and with that entry equal to 1.

Let

$$E_{ij}(c) = I + c e_i e_j^T.$$

We leave to the reader the exercise of proving the following elementary results:

Lemma 10.3.1. *The following elementary row and column operations are obtained by pre- and post-multiplying by the* elementary *matrix $E_{ij}(c)$:*

(i) $E_{ij}(c)A$ is obtained by adding c times the jth row of A to the ith row.

(ii) $AE_{ij}(c)$ is obtained by adding c times the ith column of A to the jth column.

(iii) $E_{ij}(-c)$ is the inverse of the matrix $E_{ij}(c)$.

Moreover, if T is upper triangular with diagonal $diag(\lambda_1, \ldots, \lambda_n)$, and if $\lambda_i \neq \lambda_j$ with $i < j$, then

$$T' = E_{ij}(-c)TE_{ij}(c) = E_{ij}(-c) \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & * & \\ & & \lambda_i & & & \\ & & & \ddots & & \\ & 0 & & & \lambda_j & \\ & & & & & \ddots \end{pmatrix} E_{ij}(c),$$

where T' is obtained from T by replacing the (i,j) entry t_{ij} of T with $t_{ij} + c(\lambda_i - \lambda_j)$. The only other entries of T that can possibly be affected are to the right of t_{ij} or above t_{ij}.

Using Lemma 10.3.1 over and over, starting low, working left to right in a given row and moving upward, we can transform T into a direct sum of blocks

$$P^{-1}TP = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix},$$

where each block A_i is upper triangular with each diagonal element equal to λ_i, $1 \leq i \leq r$.

Our next goal is to find an invertible matrix U_i that transforms the block A_i into a matrix $U_i^{-1}A_iU_i = \lambda_i I + N$ where N is nilpotent with a special form. All entries on or below the main diagonal are 0, all entries immediately above the diagonal are 0 or 1, and all entries further above the diagonal are 0. This matrix $\lambda_i I + N$ is called a *Jordan block*. We want to arrange it so that it is a direct sum of *elementary* Jorday blocks. These are matrices of the form $\lambda_i I + N$ where each entry N just above the diagonal is 1 and all other elements of N are 0. Then we want to use block multiplication to transform $P^{-1}TP$ into a direct sum of elementary Jordan blocks.

10.4 Transforming Nilpotent Matrices

Let B be $n \times n$, complex and nilpotent of order p: $B^{p-1} \neq 0 = B^p$. By $N(B^j)$ we mean the right null space of the matrix B^j. Note that if $j > i$, then $N(B^i) \subseteq N(B^j)$. We showed above that if $N(B^i) = N(B^{i+1})$, then $N(B^i) = N(B^k)$ for all $k \geq k$.

Step 1.

Lemma 10.4.1. *Let $W = (w_1, \ldots, w_r)$ be an independent list of vectors in $N(B^{j+1})$ with $\langle W \rangle \cap N(B^j) = \{\overline{0}\}$. Then $BW = (Bw_1, \ldots, Bw_r)$ is an independent list in $N(B^j)$ with $\langle BW \rangle \cap N(B^{j-1}) = \{\overline{0}\}$.*

Proof. Clearly $BW \subseteq N(B^j)$. Suppose $\sum_{i=1}^r c_i B w_i + u_{j-1} = \overline{0}$, with $u_{j-1} \in N(B^{j-1})$. Then $\overline{0} = B^{j-1}(\sum_{i=1}^r c_i B w_i + u_{j-1}) = \sum_{i=1}^r c_i B^j w_i + \overline{0}$. This says $\sum_{i=1}^r c_i w_i \in N(B^j)$, so by hypothesis $c_1 = c_2 = \cdots = c_r = 0$, and hence also $u_{j-1} = \overline{0}$. It is now easy to see that the Lemma must be true. □

Before proceeding to the next step, we introduce some new notation. Suppose V_1 is a subspace of V. To say that the list $(w_1, \ldots, w_r)$ is *independent in $V \setminus V_1$* means first that it is independent, and second that if W is the span $\langle w_1, \ldots, w_r \rangle$ of the given list, then $W \cap V_1 = \{\overline{0}\}$. To say that the list $(w_1, \ldots, w_r)$ is a *basis of $V \setminus V_1$* means that a basis of V_1 adjoined to the list $(w_1, \ldots, w_r)$ is a basis of V. Also, $\langle L \rangle$ denotes the space spanned by the list L.

Step 2. Let U_p be a basis of $N(B^p) \setminus N(B^{p-1}) = \mathcal{C}^n \setminus N(B^{p-1})$, since $B^p = 0$. By Lemma 10.4.1 BU_p is an independent list in $N(B^{p-1}) \setminus N(B^{p-2})$, with $\langle BU_p \rangle \cap N(B^{p-2}) = \{\overline{0}\}$.

Complete BU_p to a basis (BU_p, U_{p-1}) of $N(B^{p-1}) \setminus N(B^{p-2})$. At this point we have that

$$(BU_p, U_{p-1}, U_p) \text{ is a basis of } N(B^p) \setminus N(B^{p-2}) = \mathcal{C}^n \setminus N(B^{p-2}).$$

Step 3. (B^2U_p, BU_{p-1}) is an independent list in $N(B^{p-2}) \setminus N(B^{p-3})$. Complete this to a basis $(B^2U_p, BU_{p-1}, U_{p-2})$ of $N(B^{p-2}) \setminus N(B^{p-3})$. At this stage we have that

$$(U_p, BU_p, U_{p-1}, B^2U_p, BU_{p-1}, U_{p-2}) \text{ is a basis of } N(B^p) \setminus N(B^{p-3}).$$

Step 4. Proceed in this way until a basis for the entire space V has been obtained. At that point we will have a situation described in the following array:

$$\begin{array}{rll} \text{Independent set} & \text{basis for} & \text{subspace} \\ (U_p) & & N(B^p) \setminus N(B^{p-1}) = \mathcal{C}^n \setminus N(B^{p-1}) \\ (U_{p-1}, BU_p) & & N(B^{p-1}) \setminus N(B^{p-2}) \\ (U_{p-2}, BU_{p-1}, B^2U_p) & & N(B^{p-2}) \setminus N(B^{p-3}) \\ (B_{p-3}, BU_{p-2}, B^2U_{p-1}, B^3U_p) & & N(B^{p-3}) \setminus N(B^{p-4}) \\ \vdots & & \vdots \\ (U_1, BU_2, \ldots, B^{p-2}U_{p-1}, B^{p-1}U_p) & & N(B^{p-(p-1)}) \setminus N(B^0) = N(B) \end{array} \tag{10.11}$$

Choose $x = x_1 \in B^{p-j}U_{p-j+1}$ and follow it back up to the top of that column: $x_{p-j} \in U_{p-j}$; $Bx_{p-j} = x_{p-j-1}$; $\ldots$; $Bx_2 = x_1$. So $B^{p-j-1}x_{p-j} = x_1$.

We now want to interpret what it means for

$$P^{-1}BP = H = \begin{pmatrix} H_1 & & \\ & \ddots & \\ & & H_s \end{pmatrix} \text{ with } H_i = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & \\ & & & 1 \\ & & & 0 \end{pmatrix}.$$

This last matrix is supposed to have 1's along the superdiagonal and 0's elsewhere. This says that the jth column of BP (which is B times the jth column of P) equals the jth column of PH, which is either the zero column or the $(j-1)$st column of P. So we need

$$B(j\text{th column of } P) = \overline{0} \text{ or the } (j-1)\text{st column of } B.$$

So to form P, as the columns of P we take the vectors in the Array 10.11 starting at the bottom of a column, moving up to the top, then from the bottom to the top of the next column to the left, etc. Each column of Array 10.11 represents one block H_i, and the bottom row consists of a basis for the null space of B. Then we do have

$$P^{-1}BP = H = \begin{pmatrix} H_1 & & \\ & \ddots & \\ & & H_s \end{pmatrix}.$$

Suppose we have a matrix $A = \lambda I + B$ where B is nilpotent with all elements on or below the diagonal equal to 0. Construct P so that $P^{-1}BP = H$, i.e., $P^{-1}AP = P^{-1}(B + \lambda I)P = \lambda I + H$ is a direct sum of elementary Jordan blocks each with the same eigenvalue λ along the diagonal and 1's just above the diagonal. Such a matrix is said to be in *Jordan form.*

Start with a general $n \times n$ matrix A over C. Here is the general algorithm for finding a Jordan form for A.

1. Find P_1 so that $P_1^{-1}AP_1 = T$ is in upper triangular Schur form, so

$$T = \begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & * & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & 0 & & & & \ddots & & \\ & & & & & & & \lambda_r & \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_r \end{pmatrix}.$$

2. Find P_2 so that

$$P_2^{-1}TP_2 = \begin{pmatrix} T_{\lambda_1} & & & \\ & T_{\lambda_2} & 0 & \\ & & \ddots & \\ & 0 & & T_{\lambda_r} \end{pmatrix} = C.$$

3. Find P_3 so that $P_3^{-1}CP_3 = D + H$ where

$$(P_1P_2P_3)^{-1} \cdot A \cdot (P_1P_2P_3) = D + H \text{ is in Jordan form.}$$

An $n \times n$ *elementary Jordan block* is a matrix of the form $J_n(\lambda) = \lambda I + N_n$, where $\lambda \in \mathcal{C}$, N_n is the $n \times n$ nilpotent matrix with 1's along the superdiagonal and 0's elsewhere. The minimal and characteristic polynomials of $J_n(\lambda$ are both equal to $(x - \lambda)^n$, and J_n has a 1-dimensional space of eigenvectors spanned by $(1, 0, \ldots, 0)^T$ and an n-dimensional space of generalized eigenvectors, all belonging to the eigenvalue λ. If A has a Jordan form consisting

of a direct sum of s elementary Jordan blocks, the characteristic polynomials of these blocks are the *elementary divisors* of A. The characteristic polynomial of A is the product of its elementary divisors. If $\lambda_1, \ldots, \lambda_s$ are the distinct eigenvalues of A, and if for each i, $1 \leq i \leq s$, $(x - \lambda_i)^{m_i}$ is the largest elementary divisor involving the eigenvalue λ_i, then $\prod_{i=1}^{s}(x - \lambda_i)^{m_i}$ is the minimal polynomial for A. It is clear that $(x - \lambda)$ divides the minimal polynomial of A if and only if it divides the characteristic polynomial of A if and only if λ is an eigenvalue of A.

10.5 A "Jordan Form" for Real Matrices

Theorem 10.5.1. *Let V be a real vector space and $T \in \mathcal{L}(V)$. Then there is a basis $\mathcal{B}$ of V for which*

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

Proof. The result is clearly true if $n = \dim(V) = 1$, so suppose $n = 2$. If T has an eigenvalue λ, let v_1 be any nonzero eigenvector of T belonging to λ. Extend (v_1) to a basis (v_1, v_2) of V. With respect to this basis T has an upper triangular matrix of the form

$$\begin{pmatrix} \lambda & a \\ 0 & b \end{pmatrix}.$$

In particular, if T has an eigenvalue, then there is a basis of V with respect to which T has an upper triangular matrix. If T has no eigenvalues, then choose any basis (v_1, v_2) of V. With respect to this basis, the matrix of T has no eigenvalues. So we have the desired result when $n = 2$. Now suppose that $n = \dim(V) > 2$ and that the desired result holds for all real vector spaces with smaller dimension. If T has an eigenvalue, let U be a 1-dimensional subspace of V that is T-invariant. Otherwise, by Theorem 7.3.1 let U be a 2-dimensional T-invariant subspace of V. Choose any basis of U and let A_1 denote the matrix of $T|_U$ with respect to this basis. If A_1 is a 2-by-2 matrix then T has no eigenvalues, since otherwise we would have chosen U to be 1-dimensional. Hence $T|_U$ and A_1 have no eigenvalues.

Let W be any subspace of V for which $V = U \oplus W$. We would like to apply the induction hypothesis to the subspace W. Unfortunately, W might not be T-invariant, so we have to be a little tricky. Define $S \in \mathcal{L}(W)$ by

$$S(w) = P_{W,U}(T(w)) \ \forall w \in W.$$

Note that

$$\begin{aligned} T(w) &= P_{U,W}(T(w)) + P_{W,U}(T(w)) \\ &= P_{U,W}(T(w)) + S(w) \ \forall w \in W. \end{aligned} \tag{10.12}$$

By our induction hypothesis, there is a basis for W with respect to which S has a block upper triangular matrix of the form

$$\begin{pmatrix} S_2 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues. Adjoin this basis of W to the basis of U chosen above, getting a basis of V. The corresponding matrix of T is a block upper triangular matrix of the desired form. □

Our definition of the characteristic polynomial of the 2×2 matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ is that it is

$$f(x) = (x-a)(x-d) - bc = x^2 - (trace(A))x + det(A).$$

The fact that this is a reasonable definition (the only reasonable definition if we want the Cayley-Hamilton theorem to be valid for such matrices) follows from the next result.

Theorem 10.5.2. *Suppose V is a real vector space with dimension 2 and $T \in \mathcal{L}(V)$ has no eigenvalues. Suppose A is the matrix of T with respect to some basis. Let $p(x) \in \mathcal{R}[x]$ be a polynomial of degree 2.*

(a) If $p(x)$ equals the characteristic polynomial of A, then $p(T) = 0$.

(b) If $p(x)$ does not equal the characteristic polynomial of A, then $p(T)$ is invertible.

Proof. Part (a) follows immediately from the Cayley-Hamilton theorem, but it is also easy to derive it iindependently of that result. For part (b), let $q(x)$ denote the characteristic polynomial of A with $p(x) \neq q(x)$. Write $p(x) = x^2 + \alpha_1 x + \beta_1$ and $q(x) = x^2 + \alpha_2 x + \beta_2$ for some $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathcal{R}$. Now

$$p(T) = p(T) - q(T) = (\alpha_1 - \alpha_2)T + (\beta_1 - \beta_2)I.$$

If $\alpha_1 = \alpha_2$, then $\beta_1 \neq \beta_2$, since otherwise $p = q$. In this case $p(T)$ is some multiple of the identity and hence is invertible. If $\alpha_1 \neq \alpha_2$, then

$$p(T) = (\alpha_1 - \alpha_2)(T - \frac{\beta_2 - \beta_1)}{\alpha_1 - \alpha_2} I),$$

which is an invertible operator since T has no eigenvalues. Thus (b) holds. □

Now suppose V is 1-dimensional and $T \in \mathcal{L}(V)$. For $\lambda \in \mathcal{R}$, $null(T - \lambda I)$ equals V if λ is an eigenvalue of T and $\{\overline{0}\}$ otherwise. If $\alpha, \beta \in \mathcal{R}$ with $\alpha^2 < 4\beta$, so that $x^2 + \alpha x + \beta = 0$ has no real roots, then

$$null(T^2 + \alpha T + \beta I) = \{\overline{0}\}.$$

(Proof: Because V is 1-dimensional, there is a constant $\lambda \in \mathcal{R}$ such that $Tv = \lambda v$ for all $v \in V$. (Why is this true?) So if v is a nonzero vector in V, then $(T^2 + \alpha T + \beta I)v = (\lambda^2 + \alpha\lambda + \beta)v$. The only way this can be $\overline{0}$ is for $v = \overline{0}$ or $(\lambda^2 + \alpha\lambda + \beta) = 0$. But this second equality cannot hold because we are assuming that $\alpha^2 < 4\beta$. Hence $null(T^2 + \alpha T + \beta I) = \{\overline{0}\}$.)

Now suppose that V is a 2-dimensional real vector space and that $T \in \mathcal{L}(V)$ still has no eigenvalues. For $\lambda \in \mathcal{R}$, $null(T - \lambda I) = \{\overline{0}\}$ exactly because T has no eigenvalues. If $\alpha, \beta \in \mathcal{R}$ with $\alpha^2 < 4\beta$, then $null(T^2 + \alpha T + \beta I)$ equals all of V if $x^2 + \alpha x + \beta$ is the characteristic polynomial of T with respect to some (and hence to all) ordered bases of V, and equals $\{\overline{0}\}$ otherwise by part (b) of Theorem 10.5.2. It is important to note that in this case the null space of $T^2 + \alpha T + \beta I$ is either $\{\overline{0}\}$ or the whole space 2-dimensional space!

The goal of this section is to prove the following theorem.

Theorem 10.5.3. *Suppose that V is a real vector space of dimension n and $T \in \mathcal{L}(V)$. Suppose that with respect to some basis of V, the matrix of T has*

the form

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix}, \tag{10.13}$$

where each A_j is a 1×1 matrix or a 2×2 matrix with no eigenvalues (as in Theorem 9.4).

(a) If $\lambda \in \mathcal{R}$, then precisely dim(null$((T-\lambda I)^n)$) of the matrices $A_1, \ldots, A_m$ equal the 1×1 matrix $[\lambda]$.

(b) If $\alpha, \beta \in \mathcal{R}$ satisfy $\alpha^2 < 4\beta$, then precisely

$$\frac{dim(null(T^2 + \alpha T + \beta I)^n)}{2}$$

of the matrices $A_1, \ldots, A_m$ have characteristic polynomial equal to $x^2+\alpha x+b$.

Note that this implies that null$((T^2+\alpha T+\beta I)^n)$ must have even dimension.

Proof. We imitate Axler in constructing one proof that can be used to prove both (a) and (b). For this, let $\lambda, \alpha, \beta \in \mathcal{R}$ with $\alpha^2 < 4\beta$. Define $p(x) \in \mathcal{R}[x]$ by

$$p(x) = \begin{cases} x - \lambda, & \text{if we are trying to prove (a);} \\ x^2 + \alpha x + \beta, & \text{if we are trying to prove (b).} \end{cases}$$

Let d denote the degree of $p(x)$. Thus $d = 1$ or $d = 2$, depending on whether we are trying to prove (a) or (b).

The basic idea of the proof is to proceed by induction on m, the number of blocks along the diagonal in Eq. 10.13. If $m = 1$, then $\dim(V) = 1$ or $\dim(V) = 2$. In this case the discussion preceding this theorem implies that the desired result holds. Our induction hypothesis is that for $m > 1$, the desired result holds when m is replaced with $m - 1$.

Let $\mathcal{B}$ be a basis of V with respect to which T has the block upper-triangular matrix of Eq. 10.13. Let U_j denote the span of the basis vectors corresponding to A_j. So $\dim(U_j) = 1$ if A_j is 1×1 and $\dim(U_j) = 2$ if A_j is a 2×2 matrix (with no eigenvalues). Let

$$U = U_1 + U_2 + \cdots + U_{m-1} = U_1 \oplus U_2 \oplus \cdots \oplus U_{m-1}.$$

Clearly U is invariant under T and the matrix of $T|_U$ with respect to the basis $\mathcal{B}'$ obtained from the basis vectors corresponding to $A_1, \ldots, A_{m-1}$ is

$$[T|_U]_{\mathcal{B}'} = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_{m-1} \end{pmatrix}, \tag{10.14}$$

Suppose that $\dim(U) = n'$, so n' is either $n-1$ or $n-2$. Also, $\dim(\text{null}(p(T|_U)^{n'})) = \dim(\text{null}(p(T|_U)^n))$. Hence our induction hypothesis implies that

$$\begin{array}{l} \text{Precisely } \frac{1}{d}\dim(\text{null}(p(T|_U))^n) \text{ of the matrices} \\ A_1, \ldots, A_{m-1} \text{ have characteristic polynomial } p. \end{array} \tag{10.15}$$

Let u_m be a vector in U_m. $T(u_m)$ might not be in U_m, since the entries of the matrix A in the columns above the matrix A_m might not all be 0. But we can project $T(u_m)$ onto U_m. So let $S \in \mathcal{L}(U_m)$ be the operator whose matrix with respect to the basis corresponding to U_m is A_m. It follows that $S(u_m) = P_{U_m,U}T(u_m)$. Since $V = U \oplus U_m$, we know that for any vector $v \in V$ we have $v = P_{U,U_m}(v) + P_{U_m,U}(v)$. Putting $T(u_m)$ in place of v, we have

$$\begin{aligned} T(u_m) &= P_{U,U_m}T(u_m) + P_{U_m,U}T(u_m) \\ &= {}^*u + S(u_m), \end{aligned} \tag{10.16}$$

where ${}^*u$ denotes an unknown vector in U. (Each time the symbol ${}^*u$ is used, it denotes some vector in U, but it might be a different vector each time the symbol is used.) Since $S(u_m) \in U_m$, so $T(S(u_m)) = {}^*u + S^2(u_m)$, we can apply T to both sides of Eq. 10.16 to obtain

$$T^2(u_m) = {}^*u + S^2(u_m). \tag{10.17}$$

(It is important to keep in mind that each time the symbol ${}^*u$ is used, it probably means a different vector in U.)

Using equations Eqs. 10.16 and 10.17 it is easy to show that

$$p(T)(u_m) = {}^*u + p(S)u_m. \tag{10.18}$$

Note that $p(S)(u_m) \in U_m$. Thus iterating the last equation gives

$$p(T)^n(u_m) = {}^*u + p(S)^n(u_m). \tag{10.19}$$

Since $V = U \oplus U_m$, for any $v \in V$, we can write $v = {}^*u + u_m$, with ${}^*u \in U$ and $u_m \in U_m$. Then using Eq. 10.19 and the fact that U is invariant under any polynomial in T, we have

$$p(T)^n(v) = {}^*u + p(S)^n(u_m), \tag{10.20}$$

where $v = {}^*u + u_m$ as above. If $v \in \text{null}(p(T)^n)$, we now see that $\overline{0} = {}^*u + P(S)^n(u_m)$, from which it follows that $P(S)^n(u_m) = \overline{0}$.

The proof now breaks into two cases: Case 1 is where $p(x)$ is **not** the characteristic polynomial of A_m; Case 2 is where $p(x)$ **is** the characteristic polynomial of A_m.

So consider Case 1. Since $p(x)$ is not the characteristic polynomial of A_m, we see that $p(S)$ must be invertible. This follows from Theorem 10.5.3 and the discussion immediately following, since the dimension of U_m is at most 2. Hence $P(S)^n(u_m) = \overline{0}$ implies $u_m = \overline{0}$. This says:

$$\text{The null space of } P(T)^n \text{ is contained in } U. \tag{10.21}$$

This says that

$$\text{null}(p(T)^n) = \text{null}(p(T|_U)^n).$$

But now we can apply Eq. 10.15 to see that precisely $\left(\frac{1}{d}\right)\dim(\text{null}(p(T)^n)$ of the matrices $A_1, \ldots, A_{m-1}$ have characteristic polynmial $p(x)$. But this means that precisely $\left(\frac{1}{d}\right)\dim(\text{null}(p(T)^n)$ of the matrices $A_1, \ldots, A_m$ have characteristic polynomial $p(x)$. This completes Case 1. Now suppose that $p(x)$ is the characteristic polynomial of A_m. It is clear that $\dim(U_m) = d$.

Lemma 10.5.4. *We claim that*

$$dim(null(p(T)^n)) = dim(null(p(T|_U)^n) + d.$$

This along with the induction hypothesis Eq. 10.15 would complete the proof of the theorem.

But we still have some work to do.

Lemma 10.5.5. $V = U + null(p(T)^n)$.

Proof. Because the characteristic polynomial of the matrix A_m of S equals $p(x)$, we have $p(S) = 0$. So if $u_m \in U_m$, from Eq. 10.16 we see that $p(T)(u_m) \in U$. So

$$p(T)^n(u_m) = p(T)^{n-1}(p(T)(u_m)) \in \text{range}(p(T|_U)^{n-1}) = \text{range}(p(T|_U)^n),$$

where the last identity follows from the fact that $\dim(U) < n$. Thus we can choose $u \in U$ such that $p(T)^n(u_m) = p(T)^n(u)$. Then

$$\begin{aligned} p(T)^n(u_m - u) &= p(T)^n(u_m) - p(T)^n(u) \\ &= p(T|_U)^n(u) - p(T|_U)^n(u) \qquad (10.22) \\ &= \overline{0}. \end{aligned}$$

This says that $u_m - u \in \text{null}(p(T)^n)$. Hence u_m, which equals $u + (u_m - u)$, is in $U + \text{null}(p(T)^n)$, implying $U_m \subseteq U + \text{null}(p(t)^n)$. Therefore $V = U + \text{null}(p(T)^n) \subseteq V$. This proves the lemma 10.5.5. □

Since $\dim(U_m) = d$ and $\dim(U) = n - d$, we have

$$\begin{aligned} \dim(\text{null}(p(T)^n) &= \dim(U \cap \text{null}(p(T)^n) + \dim(U + \text{null}(p(T)^n) - \dim U \\ &= \dim(\text{null}(p(T|_U)^n)) + \dim(V) - (n-d) \qquad (10.23) \\ &= \dim(\text{null}(p(T|_U)^n)) + d. \qquad (10.24) \end{aligned}$$

This completes a proof of the claim in Lemma 10.5.4, and hence a proof of Theorem 10.5.3 □

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. An ordered pair (α, β) of real numbers is called an *eigenpair* of T if $\alpha^2 < 4\beta$ and

$$T^2 + \alpha T + \beta I$$

is not injective. The previous theorem shows that T can have only finitely many eigenpairs, because each eigenpair correpsponds to the characteristic polynomial of a 2×2 matrix on the diagonal of A in Eq. 10.13, and there is room for only finitely many such matrices along that diagonal.

We define the *multiplicity* of an eigenpair (α, β) of T to be

$$\frac{\dim(\text{null}(T^2 + \alpha T + \beta I)^{\dim(V)})}{2}.$$

From Theorem 10.5.3 we see that the multiplicity of (α, β) equals the number of times that $x^2 + \alpha x + \beta$ is the characteristic polynomial of a 2×2 matrix on the diagonal of A in Eq. 10.14.

Theorem 10.5.6. *If V is a real vector space and $T \in \mathcal{L}(V)$, then the sum of the multiplicities of all the eigenvalues of T plus the sum of twice the multiplicities of all the eigenpairs of T equals dim(V).*

Proof. There is a basis of V with respect to which the matrix of T is as in Theorem 10.5.3. The multiplicity of an eigenvalue λ equals the number of times the 1×1 matrix $[\lambda]$ appears on the diagonal of this matrix (from 10.5.3). The multiplicity of an eigenpair (α, β) equal the number of times $x^2 + \alpha x + \beta$ is the characteristic polynomial of a 2×2 matirx on the diagonal of this matrix (from 10.5.3). Because the diagonal of this matrix has length $\dim(V)$, the sum of the multiplicities of all the eigenvalues of T plus the sum of twice the multiplicities of all the eigenpairs of T must equal $\dim(V)$. □

Axler's approach to the characteristic polynomial of a real matrix is to define them for matrices of sizes 1 and 2, and then define the characteristic polynomial of a real matrix A as follows. First find the matrix $J = P^{-1}AP$ that is the Jordan form of A. Then the characteristic polynomial of A is the product of the characteristic polynomials of the 1×1 and 2×2 matrices along the diagonal of J. Then he gives a fairly involved proof (page 207) that the Cayley-Hamilton theorem holds, i.e., the characteristic polynomial of a matrix A has A as a zero. So the minimal polynomial of A divides the characteristic polynomial of A.

Theorem 10.5.7. *Suppose V is a real, n-dimensional vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with $U_1, \ldots, U_m$ the corresponding spaces of generalized eigenvectors. So $U_j = null(T - \lambda_j I)^n$. Let $(\alpha_1, \beta_1), \ldots, (\alpha_r, \beta_r)$ be the distinct eigenpairs of T. Let $V_j = null(T^2 + \alpha_j T + \beta_j)^n$, for $1 \leq j \leq r$. Then*

(a) $V = U_1 \oplus \cdots \oplus U_m \oplus V_1 \oplus \cdots \oplus V_r$.

(b) Each U_i and each V_j are invariant under T.

(c) Each $(T - \lambda_i I)|_{U_i}$ and each $(T^2 + \alpha_j T + \beta_j I)|_{V_j}$ are nilpotent.

10.6 Exercises

1. Suppose A is a block upper-triangular matrix

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_j is a square matrix. Prove that the set of eigenvalues of A equals the union of the sets of eigenvalues of $A_1, \ldots, A_m$.

Solution: We first prove that 0 is an eigenvalue of A if and only if 0 is an eigenvalue of at least one of the A_k's. So suppose A is $n \times n$ and each A_j is $n_j \times n_j$. Write a typical $n \times 1$ matrix X in the form

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix},$$

where each x_j is $n_j \times 1$. The product AX can be computed using block multiplication. First suppose that 0 is an eigenvalue of A. Thus there exists a nonzero $n \times 1$ matrix X such that $AX = \overline{0}$. Write X in the form given above and let k be the largest index with nonzero x_k. Thus

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

(If $k = m$, then the 0's shown above at the tail of X do not appear.) If we break AX into blocks of the same size as was done for X, then the k^{th} block of AX will equal $A_k x_k$. This follows from the block upper-triangular form of A and the 0's that appear in X after the k^{th} block. But $AX = \overline{0}$, so the k^{th} block of AX equal $\overline{0}$, so $A_k x_k = \overline{0}$. Because $x_k \neq \overline{0}$, this implies that 0 is an eigenvalue of A_k, as desired.

For the converse, suppose that 0 is an eigenvalue of some A_k. This means that the operator on $M_{n_k,1}(F)$ that sends $x_k \in M_{n_k,1}(F)$ to $A_k x_k$ is not injective, and hence is not surjective. Then the operator T_k on $M_{n_1+n_2+\cdots+n_k,1}(F)$ that sends

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$$

to

$$\begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & a_k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$$

is not surjective (because the last block in the product above will be $A_k x_k$, which cannot be an arbitrary $n_k \times 1$ matrix).

Again, this means that the operator T_k cannot be injective. Hence there is a nonzero vector $(x_1, \ldots, x_k)^T \in M(n_1 + \cdots + n_k, 1)(F)$ such that

$$\begin{pmatrix} A_1 & * & \\ & \ddots & \\ 0 & & A_k \end{pmatrix} \begin{pmatrix} x_1 \\ \ddots \\ x_k \end{pmatrix} = 0.$$

Now adjoinging an appropriate number of 0's gives us

$$\begin{pmatrix} A_1 & & & & & * \\ & \ddots & & & & \\ & & A_k & & & \\ & & & A_{k+1} & & \\ & & & & \ddots & \\ 0 & & & & & A_m \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

In other words, 0 is an eigenvalue of A, as desired.

So at this point we know that 0 is an eigenvalue of A if and only if it is an eigenvalue of at least one of the A_k's. Now let λ be any real number. We know that λ is an eigenvalue of A if and only if 0 is an eigenvalue

of $A - \lambda I$. So replace A by $A - \lambda I$ and replace each A_k with $A_k - \lambda I$ to complete the proof.

2. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathcal{R}$ are such that $\alpha^2 < 4\beta$. Prove that

$$\text{null}(T^2 + \alpha T + \beta I)^k$$

has even dimension for every positive integer k.

Solution: Let k be a positive integer, and let $U = \text{null}(T^2+\alpha T+\beta I)^k$. We prove that $\dim(U)$ is even.

Because U is invariant under T, we can define $S \in \mathcal{L}(U)$ by $S = T|_U$. Clearly $(S^2+\alpha S+\beta I)^k$ is the zero operator on U. Thus $S^2+\alpha S+\beta I$ is a nilpotent operator on U, which implies that $\text{null}(S^2+\alpha S+\beta I)^{\dim(U)} = U$ (by 8.8). Now part (b) of 10.5.3 applied to S and U instead of T and V, shows that $\dim(U)$ is an even integer as desired.

3. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathcal{R}$ are such that $\alpha^2 < 4\beta$ and $T^2 + \alpha T + \beta I$ is nilpotent. Prove that $\dim(V)$ is even and

$$(T^2 + \alpha T + \beta I)^{\dim(V)/2} = 0.$$

Solution: Let $S = (T^2 + \alpha T + \beta I)$. Because S is nilpotent, there is a smallest postive integer m such that $S^m = 0$. Thus

$$\{\overline{0}\} = \text{null}(S^0) \subsetneq \text{null}(S) \subsetneq \cdots \subsetneq \text{null}(S^m) = V.$$

The previous exercise states that each $\text{null}(S^k)$ has even dimension (in particular V, which equals $\text{null}(S^m)$, has even dimension). Hence the dimension must increase by at least 2 in all the proper inclusions above. Thus $\dim(V) = \dim(\text{null}(S^m))$ is at least $2m$, implying $m \leq \frac{\dim(V)}{2}$. Because $S^m = 0$, this implies that $S^{\dim(V)/2} = 0$, as desired.

Chapter 11

Matrix Functions*

11.1 Operator Norms and Matrix Norms*

Let V and W be vector spaces over F having respective norms $\|\cdot\|_V$ and $\|\cdot\|_W$. Let $T \in \mathcal{L}(V, W)$. One common problem is to understand the "size" of the linear map T in the sense of its effects on the magnitude of the inputs. For each nonzero $v \in V$, the quotient $\|T(v)\|_W / \|v\|_V$ measures the magnification caused by the transformation T on that specific vector v. An upper bound on this quotient valid for all v would thus measure the overall effect of T on the size of vectors in V. It is well known that in every finite dimensional space V, the quotient $\|T(v)\|_W / \|v\|_V$ has a maximum value which is achieved with some specific vector v_0. Note that if v is replaced by cv for some nonzero $c \in F$, the quotient is not changed. Hence if $\mathcal{O} = \{x \in V : \|x\|_V = 1\}$, then we may define a norm for T (called a *transformation norm*) by

$$\|T\|_{V,W} = \max\{\|T(v)\|_W / \|v\|_V : \vec{0} \neq v \in V\} = \max\{\|T(v)\|_W : v \in \mathcal{O}\}. \tag{11.1}$$

In this definition we could replace "max" with "supremum" in order to guarantee that $\|T\|_{V,W}$ is always defined even when V and W are not finite dimensional. However, in this text we just deal with the three specific norms we have already defined on the finite dimensional vector spaces. The norm of a linear map from V to W defined in this way is a vector norm on the vector space $\mathcal{L}(V, W)$. (See Exercise 1.) We can actually say a bit more.

Theorem 11.1.1. *Let U, V and W be vector spaces over F endowed with*

vector norms $\|\cdot\|_U$, $\|\cdot\|_V$, $\|\cdot\|_W$, *respectively. Let* $S, T \in \mathcal{L}(U,V)$ *and* $L \in \mathcal{L}(V,W)$ *and suppose they each have norms defined by Eq. 11.1 Then:*

(a) $\|T\|_{U,V} \geq 0$, *and* $\|T\|_{U,V} = 0$ *if and only if* $T(u) = \vec{0}$ *for all* $u \in V$.

(b) $\|aT\|_{U,V} = |a|\ \|T\|_{U,V}$ *for all scalars* a.

(c) $\|S+T\|_{U,V} \leq \|S\|_{U,V} + \|T\|_{U,V}$.

(d) $\|T(u)\|_V \leq \|T\|_{U,V} \cdot \|u\|_U$ *for all* $u \in U$.

(e) $\|I\|_{U,U} = 1$, *where* $I \in \mathcal{L}(U)$ *is defined by* $I(u) = u$ *for all* $u \in U$.

(f) $\|L \circ T\|_{U,W} \leq \|L\|_{V,W} \cdot \|T\|_{U,V}$.

(g) If $U = V$, *then* $\|T^i\|_{U,U} \leq (\|T\|_{U,U})^i$.

Proof. Parts (a), (d) and (e) follow immediately from the definition of the norm of a linear map. Part (b) follows easily since $|a|$ can be factored out of the appropriate maximum (or supremum). For part (c), from the triangle inequality for vector norms, we have

$$\|(S+T)(u)\|_V = \|S(u)+T(u)\|_V \leq \|S(u)\|_V + \|T(u)\|_V \leq (\|S\|_{U,V} + \|T\|_{U,V})\|u\|_U,$$

from part (d). This says the quotient used to define the norm of $S+T$ is bounded above by $(\|S\|_{U,V} + \|T\|_{U,V})$, so the least upper bound is less than or equal to this. Part (f) follows in a similar manner, and then (g) follows by repeated applications of part (f). □

Let A be an $m \times n$ matrix over F (with F a subfield of $\mathcal{C}$). As usual, we may consider the linear map $T_A : F^n \to F^m : x \mapsto Ax$. We consider the norm on $\mathcal{L}(F^n, F^m)$ induced by each of the standard norms $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$, and use it to define a corresponding norm on the vector space of $m \times n$ matrices. It seems natural to let the transformation norm of T_A also be a "norm" of A. Specifically, we have

$$\begin{aligned} \|A\|_1 &= \max\left\{\tfrac{\|Ax\|_1}{\|x\|_1} : x \neq \vec{0}\right\} \\ \|A\|_2 &= \max\left\{\tfrac{\|Ax\|_2}{\|x\|_2} : x \neq \vec{0}\right\} \\ \|A\|_\infty &= \max\left\{\tfrac{\|Ax\|_\infty}{\|x\|_\infty} : x \neq \vec{0}\right\}. \end{aligned}$$

Theorem 11.1.2. *Let* A *be* $m \times n$ *over the field* F. *Then:*

(a) $\|A\|_1 = max\{\sum_{i=1}^m |A_{ij}| : 1 \leq j \leq n\}$ *(maximum absolute column sum).*

(b) $\|A\|_\infty = max\{\sum_{j=1}^n |A_{ij}| : 1 \leq i \leq m\}$ *(maximum absolute row sum).*

(c) $\|A\|_2 =$ *maximum singular value of* A.

Proof. To prove (a), observe

$$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} A_{ij} x_j \right| \leq \sum_{i=1}^{m} \sum_{j=1}^{n} |A_{ij}| \ |x_j|$$

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} |A_{ij}| \right) |x_j| \leq \sum_{j=1}^{n} \left(\max_j \ \sum_i |A_{ij}| \right) |x_j|$$

$$= \left(\max_j \sum_{i=1}^{m} |A_{ij}| \right) \|x\|_1 = \alpha \|x\|_1.$$

So $\|A\|_1 \leq \alpha = \max_j \sum_{i=1}^{m} |A_{ij}|$. To complete the proof of (a) we need to construct a vector $x \in F^n$ such that $\|Ax\|_1 = \alpha \|x\|_1$. Put $x = (0, 0, \ldots, 1, \ldots, 0)^T$ where the single nonzero entry 1 is in column j_0, and the maximum value of $\sum_{i=1}^{m} |A_{ij}|$ occurs when $j = j_0$. Then $\|x\|_1 = 1$, and $\|Ax\|_1 = \sum_{i=1}^{m} |A_{ij_0}| = \alpha = \alpha \|x\|_1$ as desired.

We now consider part (b). Define α by

$$\alpha = \max\{\sum_{j=1}^{n} |A_{ij}| : 1 \leq i \leq m\}.$$

For an arbitrary $v = (v_1, \ldots, v_n)^T \in F^n$ suppose that the maximum defining α above is attained when $i = i_0$. Then we have

$$\|T_A(v)\|_\infty = \|Av\|_\infty = \max_i |(Av)_i| = \max_i \left| \sum_{j=1}^{n} A_{ij} v_j \right|$$

$$\leq \max_i \sum_{j=1}^{n} (|A_{ij}| \ |v_j|) \leq \max_i \sum_{j=1}^{n} (|A_{ij}| \ \max_k |v_k|)$$

$$\leq \alpha \|v\|_\infty,$$

so $\|T_A(v)\|_\infty / \|v\|_\infty \leq \alpha$ for all nonzero v. This says that this norm of T_A is at most α. To show that it equals α we must find a vector v such that the usual quotient equals α. For each $j = 1, \ldots, n$, chose v_j with absolute value 1 so that $A_{i_0 j} v_j = |A_{i_0 j}|$. Clearly $\|v\|_\infty = 1$, so that $\|T_A(v)\|_\infty \leq \alpha$, which we knew would have to be the case anyway. On the other hand

$$|(T_A(v))_{i_0}| = \left|\sum_{j=1}^{n} A_{i_0 j} v_j\right| = \left|\sum_{j=1}^{n} |A_{i_0 j}|\right| = \alpha.$$

This shows that $\|T_A(v)\|_\infty = \alpha$ for this particular v, and hence this norm of T_A is α.

Part (c) is a bit more involved. First note that $\|v\|$ as used in earlier chapters is just $\|v\|_2$. Suppose that S is unitary and $m \times m$, and that A is $m \times n$. Then $\|(SA)(v)\|_2 = \|S(A(v))\|_2 = \|(A(v)\|_2$ for all $v \in F^n$. It follows immediately that $\|SA\|_2 = \|A\|_2$. Now let A be an arbitrary $m \times n$ matrix with singular value decomposition $A = U\Sigma V^*$. So $\|A\|_2 = \|U\Sigma V^*\|_2 = \|\Sigma V^*\|_2$. Since $\|\cdot\|_2$ is a matrix norm coming from a transformation norm, by part (f) of Theorem 11.1.1 $\|\Sigma V^*\|_2 \leq \|\Sigma\|_2 \, \|V^*\|_2 = \|\Sigma\|_2$. If $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, 0, \ldots, 0)$, with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$, then $\|\Sigma x\|_2 = \|(\sigma_1 x_1, \ldots, \sigma_k x_k, 0, \ldots, 0)^T\|_2 = \sqrt{\sigma_1^2 x_1^2 + \cdots \sigma_k^2 x_k^2} \leq \sigma_1 \cdot \|x\|_2$. So $\|\Sigma\|_2 \leq \sigma_1$. Suppose $x = (1, 0, \ldots, 0)^T$. Then $\Sigma x = (\sigma_1 x_1, 0, \ldots, 0)^T$. So $\|x\|_2 = 1$, $\|\Sigma x\|_2 = |\sigma_1 x_1| = \sigma_1 = \sigma_1 \|x\|_2$. This says $\|\Sigma\|_2 = \sigma_1$. So we know that $\|\Sigma V^*\|_2 \leq \sigma_1$. Let x be the first column of V, so $V^* x = (1, 0, \ldots, 0)^T$. Then $\|\Sigma V^* x\|_2 = \sigma_1$, showing that $\|\Sigma V^*\|_2 = \|\Sigma\|_2 = \sigma_1$, so that finally we see $\|A\|_2 = \sigma_1$. □

11.2 Polynomials in an Elementary Jordan Matrix*

Let N_n denote the $n \times n$ matrix with all entries equal to 0 except for those along the super diagonal just above the main diagonal:

$$N_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}. \tag{11.2}$$

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

The following lemma is easily established

Lemma 11.2.1. *For* $1 \leq m \leq n-1$, $(N_n^m)_{ij} = \begin{cases} 1, & \text{if } j = i+m; \\ 0, & \text{otherwise.} \end{cases}$ *Also,* $N_n^0 = I$ *and* $N_n^n = 0$.

Then let $J_n(\lambda)$ be the elementary $n \times n$ Jordan block with eigenvalue λ given by $J_n(\lambda) = \lambda I + N_n$. So

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} = \lambda I + N_n,$$

where N_n is nilpotent with minimal polynomial x^n.

Theorem 11.2.2. *In this Theorem (and until notified otherwise) write* J *in place of* $J_n(\lambda)$. *For each positive integer* m,

$$J^m = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \cdots & \binom{m}{m-1}\lambda^{m-n+1} \\ & \lambda^m & \cdots & \binom{m}{m-2}\lambda^{m-n+2} \\ & & \ddots & \\ & & & \lambda^m \end{pmatrix},$$

that is,

$$(J^m)_{ij} = \binom{m}{j-i}\lambda^{m-j+i} \ \forall \ 1 \leq i, j \leq n.$$

Proof. It is easy to see that the desired result holds for $m = 1$ by the definition of J, so suppose that it holds for a given $m \geq 1$. Then

$$(J^{m+1})_{ij} = (J \cdot J^m)_{ij} = \sum_{k=1}^{n} J_{ik}(J^m)_{kj} = \lambda(J^m)_{ij} + 1 \cdot (J^m)_{i+1,j}$$

$$= \lambda\binom{m}{j-i}\lambda^{m-j+i} + \binom{m}{j-i-1}\lambda^{m-j+i+1}$$

$$= \binom{m}{j-i}\lambda^{m+1-j+i} + \binom{m}{j-i-1}\lambda^{m+1-j+i} = \binom{m+1}{j-i}\lambda^{m+1-j+i}.$$

□

Now suppose that $f(x) = \sum_{m=0}^{k} a_m x^m$, so that $f(J) = a_0 I + a_1 J + \cdots + a_k J^k$. Recall (or prove by induction on s) that for $s \geq 0$,

$$f^{(s)}(x) = \sum_{m=0}^{k} a_m \binom{m}{s} s! x^{m-s}, \text{ so } \sum_{m=0}^{k} a_m \binom{m}{s} \lambda^{m-s} = \frac{f^{(s)}(\lambda)}{s!}. \tag{11.3}$$

So for $s \geq 0$,

$$(f(J))_{i,i+s} = \sum_{m=0}^{k} a_m \binom{m}{s} \lambda^{m-s} = \frac{1}{s!} f^{(s)}(\lambda).$$

This says that

$$f(J) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) & \frac{1}{2!}f^{(2)}(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ 0 & f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda) \\ 0 & 0 & f(\lambda) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda) \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & f(\lambda) \end{pmatrix}. \tag{11.4}$$

Now suppose that J is a direct sum of elementary Jordan blocks:

$$J = J_1 \oplus J_2 \oplus \cdots \oplus J_s.$$

Then for the polynomial $f(x)$ we have

$$f(J) = f(J_1) \oplus f(J_2) \oplus \cdots \oplus f(J_s).$$

Here $f(J_1), \ldots, f(J_s)$ are polynomials in separate Jordan blocks, whose values are given by Eq. 11.3. This result may be applied in the computation of $f(A)$ even when A is not originally in Jordan form. First determine a T for which $J = T^{-1}AT$ has Jordan form. Then compute $f(J)$ as above and use

$$f(A) = f(TJT^{-1}) = Tf(J)T^{-1}.$$

11.3 Scalar Functions of a Matrix*

For certain functions $f : F \to F$ (satisfying requirements stipulated below) we can define a matrix function $f : A \mapsto f(A)$ so that if f is a polynomial, the

value $f(A)$ agrees with the value given just above. Start with an arbitrary square matrix A over F and let $\lambda_1, \ldots, \lambda_s$ be the distinct eigenvalues of A. Reduce A to Jordan form

$$T^{-1}AT = J_1 \oplus J_2 \oplus \cdots \oplus J_t,$$

where $J_1, \ldots, J_t$ are elementary Jordan blocks. Consider the Jordan block

$$J_i = J_{n_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ & \lambda_i & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ & & & & \lambda_i \end{pmatrix}, \tag{11.5}$$

which has $(x - \lambda_i)^{n_i}$ as its minimal and characteristic polynomials. If the function $f : F \to F$ is defined in a neighborhood of the point λ_i and has derivatives $f^{(1)}(\lambda_i), \ldots, f^{(n_i-1)}(\lambda_i)$, then define $f(J_i)$ by

$$f(J_i) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) & \frac{1}{2!}f^{(2)}(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ 0 & f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda) \\ 0 & 0 & f(\lambda) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda) \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & f(\lambda) \end{pmatrix}. \tag{11.6}$$

This says that

$$(f(J_i))_{r,r+j} = \frac{1}{j!}f^{(j)}(\lambda), \quad \text{for } 0 \le j \le n - r. \tag{11.7}$$

If f is defined in a neighborhood of each of the eigenvalues $\lambda_1, \ldots, \lambda_s$ and has (finite) derivatives of the proper orders in these neighborhoods, then also

$$f(J) = f(J_1) \oplus f(J_2) \oplus \cdots \oplus f(J_t), \tag{11.8}$$

and

$$f(A) = Tf(J)T^{-1} = Tf(J_1)T^{-1} \oplus \cdots \oplus Tf(J_t)T^{-1}. \tag{11.9}$$

The matrix $f(A)$ is called the *value* of the function f at the matrix A. We will show below that $f(A)$ does not depend on the method of reducing A to Jordan form (i.e., it does not depend on the particular choice of T),

and thus f really defines a function on the $n \times n$ matrices A. This matrix function is called the *correspondent* of the *numerical* function f. Not all matrix functions have corresponding numerical functions. Those that do are called *scalar functions*.

Here are some of the simplest properties of scalar functions.

Theorem 11.3.1. *It is clear that the definition of scalar function was chosen precisely so that part (a) below would be true.*

(a) If $f(\lambda)$ is a polynomial in λ, then the value $f(A)$ of the scalar functioin f coincides with the value of the polynomial $f(\lambda)$ evaluated at $\lambda = A$.

(b) Let A be a square matrix over F and suppose that $f_1(\lambda)$ and $f_2(\lambda)$ are two numerical functions for which the expressions $f_1(A)$ and $f_2(A)$ are meaningful. If $f(\lambda) = f_1(\lambda) + f_2(\lambda)$, then $f(A)$ is also meaningful and $f(A) = f_1(A) + f_2(A)$.

(c) With A, f_1 and f_2 as in the preceding part, if $f(\lambda) = f_1(\lambda)f_2(\lambda)$, then $f(A)$ is meaningful and $f(A) = f_1(A)f_2(A)$.

(d) Let A be a matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, each appearing as often as its algebraic multiplicity as an eigenvalue of A. If $f : F \to F$ is a numerical function and $f(A)$ is defined, then the eigenvalues of the matrix $f(A)$ are $f(\lambda_1), \ldots, f(\lambda_n)$.

Proof. The proofs of parts (b) and (c) are analogous so we just give the details for part (c). To compute the values $f_1(A)$, $f_2(A)$ and $f(A)$ according to the definition, we must reduce A to Jordan form J and apply the formulas given in Eqs. 11.8 and 11.9. If we show that $f(J) = f_1(J)f_2(J)$, then from Eq. 11.8 we immediately obtain $f(A) = f_1(A)f_2(A)$. In fact,

$$f(J) = f(J_1) \oplus \cdots \oplus f(J_t),$$

and

$$f_1(J)f_2(J) = f_1(J)f_2(J) = f_1(J_1)f_2(J_1) \oplus \cdots \oplus f_1(J_t)f_2(J_t),$$

so that the proof is reduced to showing that

$$f(J_i) = f_1(J_i)f_2(J_i), \quad (i = 1, 2, \ldots, t),$$

where J_i is an elementary Jordan block.

Start with the values of $f_1(J_i)$ and $f_2(J_i)$ given in Eq. 11.6, and multiply them together to find that

$$[f_1(J_i)f_2(J_i)]_{r,r+s} = \sum_{j=0}^{s} (f_1(J_i))_{r,r+j}(f_2(J_i))_{r+j,r+s} =$$

$$= \sum_{j=0}^{s} \frac{1}{j!} f_1^{(j)}(\lambda_i) \frac{1}{(s-j)!} f_2^{(s-j)}(\lambda_i) =$$

$$= \frac{1}{s!} \sum_{j=0}^{s} \frac{s!}{j!(s-j)!} f_1^{(j)}(\lambda_i) f_2^{(s-j)}(\lambda_i) = \frac{1}{s!}(f_1 f_2)^{(s)}(\lambda_i),$$

where the last equality comes from Exercise 2, part (i).

Thus $f_1(J_i)f_2(J_i) = f(J_i)$, completing the proof of part (c) of the theorem. For part (d), the eigenvalues of the matrices $f(A)$ and $T^{-1}f(A)T = f(T^{-1}AT)$ are equal, and therefore we may assume that A has Jordan form. Formulas in Eq. 11.5 and 11.6 show that in this case $f(A)$ is upper triangular with $f(\lambda_1), \ldots, f(\lambda_n)$ along the main diagonal. Since the diagonal elements of an upper triangular matrix are its eigenvalues, part (d) is proved. □

Similarly, if f and g are functions such that $f(g(A))$ is defined, and if $h(\lambda) = f(g(\lambda))$, then $h(A) = f(g(A))$.

To finish this section we consider two examples.

Example 11.3.2. *Let $f(\lambda) = \lambda^{-1}$. This function is defined everywhere except at $\lambda = 0$, and has finite derivatives of all orders everywhere it is defined. Consequently, if the matrix A does not have zero as an eigenvalue, i.e., if A is nonsingular, then $f(A)$ is meaningful. But $\lambda \cdot f(\lambda) = 1$, hence $A \cdot f(A) = I$, so $f(A) = A^{-1}$. Thus the matrix function $A \mapsto A^{-1}$ corresponds to the numerical function $\lambda \mapsto \lambda^{-1}$, as one would hope.*

Example 11.3.3. *Let $f(\lambda) = \sqrt{\lambda}$. To remove the two-valuedness of $\sqrt{\cdot}$ it is sufficient to slit the complex plane from the origin along a ray not containing any eigenvalues of A, and to consider one branch of the radical. Then this function, for $\lambda \neq 0$, has finite derivatives of all orders. It follows that the expression $\sqrt{A}$ is meaningful for all nonsingular matrices A. Putting $\lambda = A$ in the equation*

$$f(\lambda)f(\lambda) = \lambda,$$

we obtain

$$f(A)f(A) = A.$$

This shows that each nonsingular matrix has a square root.

11.4 Scalar Functions as Polynomials*

At this point we need to generalize the construction given in Lagrange interpolation (see Section 5.3).

Lemma 11.4.1. *Let $r_1, \ldots, r_s$ be distinct complex numbers, and for each i, $1 \leq i \leq s$, let m_i be a nonnegative integer. Let (a_{ij}) be a table of arbitrary numbers , $1 \leq i \leq s$, and for each fixed i, $0 \leq j \leq m_i$. Then there exists a polynomial $p(x)$ such that $p^{(j)}(r_i) = a_{ij}$, for $1 \leq i \leq s$ and for each fixed i, $0 \leq j \leq m_i$.*

Proof. It is convenient first to construct auxiliary polynomials $p_i(x)$ such that $p_i(x)$ and its derivatives to the m_ith order assume the required values at the point r_i and are all zero at the other given points. Put

$$\phi_i(x) = b_{i0} + b_{i1}(x - r_i) + \cdots + b_{im_i}(x - r_i)^{m_i} = \sum_{j=0}^{m_i} b_{ij}(x - r_i)^j,$$

where the b_{ij} are complex numbers to be determined later. Note that $\phi_i^{(j)}(r_i) = j!b_{ij}$.

Set

$$\Phi_i(x) = (x - r_1)^{m_i+1} \cdots (x - r_{i-1})^{m_i+1}(x - r_{i+1})^{m_i+1} \cdots (x - r_s)^{m_i+1},$$

and

$$p_i(x) = \phi_i(x)\Phi_i(x) = \left\{ \sum_{j=0}^{m_i} b_{ij}(x - r_i)^j \right\} \prod_{j \neq i}^{1 \leq j \leq s} (x - r_j)^{m_i+1}.$$

By the rule for differentiating a product (see Exercise 2),

$$p_i^{(j)}(r_i) = \sum_{l=0}^{j} \binom{j}{l} \phi_i^{(l)}(r_i)\Phi_i^{(j-l)}(r_i),$$

or

$$a_{ij} = \sum_{l=0}^{j} \binom{j}{l} l! b_{il} \Phi_i^{(j-l)}(r_i). \tag{11.10}$$

Using Eq. 11.10 with $j = 0$ and the fact that $\Phi_i(r_i) \neq 0$ we find

$$b_{i0} = \frac{a_{i0}}{\Phi_i(r_i)}, \quad 1 \leq i \leq s. \tag{11.11}$$

For each $i = 1, 2, \ldots, s$, and for a given j with $0 \leq j < m_i$, once b_{il} is determined for all l with $0 \leq l < j$, we can solve Eq. 11.10 for

$$b_{ij} = \frac{a_{ij}}{j!\Phi_i(r_i)} - \frac{\sum_{l=0}^{j-1} b_{il} \Phi_i^{(j-l)}(r_i)}{(j-l)!\Phi_i(r_i)}, \ 1 \leq i \leq s. \tag{11.12}$$

This determines $p_i(x)$ so that $p_i(x)$ and its derivatives up to the m_ith derivative have all the required values at r_i and all equal zero at r_t for $t \neq i$. It is now clear that the polynomial $p(x) = p_1(x) + p_2(x) + \cdots + p_s(x)$ satisfies all the requirements of the Lemma. □

Consider a numerical function $\lambda \mapsto f(\lambda)$ and an $n \times n$ matrix A for which the value $f(A)$ is defined. We show that there is a polynomial $p(x)$ for which $p(A)$ equals $f(A)$. Let $\lambda_1, \ldots, \lambda_s$ denote the distinct eigenvalues of the matrix A. Using only the proof of the lemma we can construct a polynomial $p(x)$ which satisfies the conditions

$$p(\lambda_i) = f(\lambda_i), \ p'(\lambda_i) = f'(\lambda_i), \ldots, p^{(n-1)}(\lambda_i) = f^{(n-1)}(\lambda_i), \tag{11.13}$$

where if some of the derivatives $f^{(j)}(r_i)$ are superfluous for the determination of $f(A)$, then the corresponding numbers in Eq. 11.12 may be replaced by zeros. Since the values of $p(x)$ and $f(\lambda)$ (and their derivatives) coincide at the numbers λ_i, then $f(A) = p(A)$. This completes a proof of the following theorem.

Theorem 11.4.2. *The values of all scalar functions in a matrix A can be expressed by polynomials in A.*

Caution: The value $f(A)$ of a given scalar function f can be represented in the form of some polynomial $p(A)$ in A. However, this polynomial, for a given function f will be different for different matrices A. For example, the

minimal polynomial $p(x)$ of a nonsingular matrix A can be used to write A^{-1} as a polynomial in A, but for different matrices A different polynomials will occur giving A^{-1} as a polynomial in A.

Also, considering the function $f(\lambda) = \sqrt{\lambda}$, we see that *for every nonsingular matrix A there exists a polynomial $p(x)$ for which*

$$p(A)p(A) = A.$$

VERY IMPORTANT: With the help of Theorem 11.4.2 we can now resolve the question left open just preceding Theorem 11.3.1 concerning whether $f(A)$ was well-defined. If we know the function $f(\lambda)$ and its derivatives at the points $\lambda_1, \ldots, \lambda_s$, we can construct the polynomial $p(x)$ whose value $p(A)$ does not depend on the reduction of the matrix A to Jordan form, and at the same time is equal to $f(A)$. Consequently, the value $f(A)$ defined in the preceding section using the reduction of A to Jordan form, does not depend on the way this reduction is carried out.

Let $f(\lambda)$ be a numerical function, and let A be a matrix for which $f(A)$ is meaningful. By Theorem 11.4.2 we can find a polynomial $p(x)$ for which $p(A) = f(A)$. For a given function $f(\lambda)$, the polynomial $p(x)$ depends only on the elementary divisors of the matrix A. But the elementary divisors of A and its transpose A^T coincide, so $p(A^T) = f(A^T)$. Since for a polynomial $p(x)$ we have $p(A^T) = p(A)^T$, it must be that $f(A^T) = f(A)^T$ for all scalar functions $f(\lambda)$ for which $f(A)$ is meaningful.

11.5 Power Series*

All matrices are $n \times n$ over F, as usual. Let

$$G(x) = \sum_{k=0}^{\infty} a_k x^k$$

be a power series with coefficients from F. Let $G_N(x) = \sum_{k=0}^{N} a_k x^k$ be the N^{th} partial sum of $G(x)$. For each $A \in M_n(F)$ let $G_N(A)$ be the element of $M_n(F)$ obtained by substituting A iun this polynomial. For each fixed i, j we obtain a sequence of real or complex numbers c_{ij}^N, $N = 0, 1, 2, \ldots$ by taking c_{ij}^N to be the (i, j) entry of the matrix $G_N(A)$. The series

$$G(A) = \sum_{k=0}^{\infty} a_k A^k$$

is said to *converge to the matrix* C in $M_n(F)$ if for each $i, j \in \{1, 2, \ldots, n\}$ the sequence $\{c_{ij}^N\}_{N=0}^{\infty}$ converges to the (i, j) entry of C (in which case we write $G(A) = C$). We say $G(A)$ *converges* if there is some $C \in M_n(F)$ such that $G(A) = C$.

For $A = (a_{ij}) \in M_n(F)$ define

$$\|A\| = \sum_{i,j=1}^{n} |a_{ij}|,$$

i.e., $\|A\|$ is the sum of the absolute values of all the entries of A.

Lemma 11.5.1. *For all $A, B \in M_n(F)$ and all $a \in F$*
(a) $\|A + B\| \leq \|A\| + \|B\|$;
(b) $\|AB\| \leq \|A\| \cdot \|B\|$;
(c) $\|aA\| = |a| \cdot \|A\|$.

Proof. The proofs are rather easy. We just give a proof of (b).

$$\|AB\| = \sum_{i,j} |(AB)_{ij}| = \sum_{i,j} \left| \sum_{k} A_{ik} B_{kj} \right| \leq \sum_{i,j,k} |A_{ik}| \cdot |B_{kj}| \leq$$

$$\leq \sum_{i,j,k,r} |A_{ik}| \cdot |B_{rj}| = \|A\| \cdot \|B\|.$$

□

Suppose that $G(x) = \sum_{k=0}^{\infty} a_k x^k$ has radius of convergence equal to R. Hence if $\|A\| < R$, then $\sum_{k=0}^{\infty} a_k \|A\|^k$ converges absolutely, i.e., $\sum_{k=0}^{\infty} |a_k| \cdot \|A\|^k$ converges. But then $|a_k (A^K)_{ij}| \leq \left| a_k \|A^k\| \right| \leq |a_k| \cdot \|A\|^k$, implying that $\sum_{k=0}^{\infty} |a_k (A^k)_{ij}|$ converges, hence $\sum_{k=0}^{\infty} a_k (A^k)_{ij}$ converges. At this point we have shown the following:

Lemma 11.5.2. *If $\|A\| < R$ where R is the radius of convergence of $G(x)$, then $G(A)$ converges to a matrix C.*

We now give an example of special interest. Let $f(\lambda)$ be the numerical function

$$f(\lambda) = exp(\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}.$$

Since this power series converges for all complex numbers λ, each matrix $A \in M_n(F)$ satisfies the condition in Lemma 11.5.2 so $\exp(A)$ converges to a matrix $C = e^A$. Also, we know that $f(\lambda) = exp(\lambda)$ has derivatives of all orders at each complex number. Hence $f(A)$ is meaningful in the sense of the preceding section. We want to be sure that the matrix C to which the series $\sum_{k=0}^{\infty} \frac{1}{k!}A^k$ converges is the same as the value $f(A)$ given in the preceding section. So let us start this section over.

A sequence of square matrices

$$A_1, A_2, \cdots, A_m, A_{m+1}, \cdots, \tag{11.14}$$

all of the same order, is said to *converge* to the matrix A provided the elements of the matrices in a fixed row and column converge to the corresponding element of the matrix A.

It is clear that if the sequences $\{A_m\}$ and $\{B_m\}$ converge to matrices A and B, respectively, then $\{A_m + B_m\}$ and $\{A_m B_m\}$ converge to $A + B$ and AB, respectively. In particular, if T is a constant matrix, and the sequence $\{A_m\}$ converges to A, then the sequence $\{T^{-1}A_m T\}$ will converge to $T^{-1}AT$. Further, if

$$A_m = A_m^{(1)} \oplus \cdots \oplus A_m^{(s)}, \ (m = 1, 2, \ldots),$$

where the orders of the blocks do not depend on m, then $\{A_m\}$ will converge to some limit if and only if each block $\{A_m^{(i)}\}$ converges separately.

The last remark permits a completely simple solution of the question of the convergence of a matrix power series. Let

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m + \cdots \tag{11.15}$$

be a formal power series in an indeterminate x. The expression

$$a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m + \cdots \tag{11.16}$$

is called the corresponding *power series* in the matrix A, and the polynomial

$$f_n(A) = a_0 I + a_1 A + \cdots + a_n A^n$$

is the *nth partial sum* of the series. The series in Eq. 11.16 is *convergent* provided the sequence $\{f_n(A)\}_{n=1}^{\infty}$ of partial sums has a limit. If this limit exists it is called the *sum* of the series.

Reduce the matrix A to Jordan form:

$$T^{-1}AT = J = J_1 \oplus \cdots \oplus J_t,$$

where $J_1, \cdots, J_t$ are elementary Jordan blocks. We have seen above that convergence of the sequence $\{f_n(A)\}$ is equivalent to the convergence of the sequence $\{T^{-1}f_n(A)T\}$. But

$$T^{-1}f_n(A)T = f_n(T^{-1}AT) = f_n(J) = f_n(J_1) \oplus \cdots \oplus f_n(J_t),$$

and the question of the convergence of the series Eq. 11.16 is equivalent to the following: Under what conditions is this series convergent for the Jordan blocks $J_1, \cdots, J_t$? Consider one of these blocks, say J_i. Let it have the elementary divisor $(x - \lambda_i)^{n_i}$, i.e., it is an elementary Jordan block with minimal and characteristic polynomial equal to $(x - \lambda_i)^{n_i}$. By Eq. 11.4

$$f_n(J_i) = \begin{pmatrix} f_n(\lambda_i) & \frac{1}{1!}f_n^{(1)}(\lambda_i) & \frac{1}{2!}f_n^{(2)}(\lambda_i) & \cdots & \frac{1}{(n-1)!}f_n^{(n-1)}(\lambda_i) \\ 0 & f_n(\lambda_i) & \frac{1}{1!}f_n^{(1)}(\lambda_i) & \cdots & \frac{1}{(n-2)!}f_n^{(n-2)}(\lambda_i) \\ 0 & 0 & f_n(\lambda_i) & \cdots & \frac{1}{(n-3)!}f_n^{(n-3)}(\lambda_i) \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & f(\lambda) \end{pmatrix}. \quad (11.17)$$

Consequently, $\{f_n(J_i)\}$ converges if and only if the sequences $\{f_n^{(j)}(\lambda_i)\}_{n=1}^{\infty}$ for each $j = 0, 1, \ldots, n_i - 1$ converge, i.e., if and only if the series Eq. 11.15 converges, and the series obtained by differentiating it term by term up to $n_i - 1$ times, inclusive, converges. It is known from the theory of analytic functions that all these series are convergent if either λ_i lies inside the circle of convergence of Eq. 11.15 or λ_i lies on the circle of convergence and the $(n_i - 1)$st derivative of Eq. 11.15 converges at λ_i. Moreover, when λ_i lies inside the circle of convergence of Eq. 11.15, then the derivative of the series evaluated at λ_i gives the derivative of the original function evaluated at λ_i. Thus we have

Theorem 11.5.3. *A matrix power series in A converges if and only if each eigenvalue λ_i of A either lies inside the circle of convergence of the corresponding power series $f(\lambda)$ or lies on the circle of convergence, and at the same time the series of $(n_i - 1)$st derivatives of the terms of $f(\lambda)$ converges at λ_i to the derivative of the function given by the original power series evaluated at λ_i, where n_i is the highest degree of an elementary divisor belonging*

to λ_i (i.e., where n_i is the size of the largest elementary Jordan block with eigenvalue λ_i). Moreover, if each eigenvalue λ_i of A lies inside the circle of convergence of the power series $f(\lambda)$, then the jth derivative of the power series in A converges to the jth derivative of $f(\lambda)$ evaluated at A.

Now reconsider the exponential function mentioned above. We know that $f(\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$ converges, say to e^λ, for all complex numbers λ. Moreover, the function $f(\lambda) = e^\lambda$ has derivatives of all orders at each $\lambda \in \mathcal{C}$ and the power series obtained from the original by differentiating term by term converges to the derivative of the function at each complex number λ. In fact, the derivative of the function is again the original function, and the power series obtained by differentiating the original power series term by term is again the original power series. Hence the power series $\sum_{k=0}^{\infty} \frac{1}{k!}A^k$ not only converges to some matrix C, it converges to the value $\exp(A)$ defined in the previous section using the Jordan form of A. It follows that for each complex $n \times n$ matrix A there is a polynomial $p(x)$ such that $\exp(A) = p(A)$.

Let J_λ denote the $n \times n$ elementary Jordan block

$$J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \lambda & 1 \\ 0 & \dots & \dots & \dots & \lambda \end{pmatrix} = \lambda I + N_n.$$

If $f(\lambda) = e^\lambda$, then we know

$$f(J_\lambda) = \begin{pmatrix} e^\lambda & \frac{1}{1!}e^\lambda & \frac{1}{2!}e^\lambda & \dots & \frac{1}{(n-1)!}e^\lambda \\ 0 & e^\lambda & \frac{1}{1!}e^\lambda & \dots & \frac{1}{(n-2)!}e^\lambda \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & e^\lambda & \frac{1}{1!}e^\lambda \end{pmatrix}.$$

Now let t be any complex number. Let $B_n(t) = B_t$ be the $n \times n$ matrix defined by

$$B_t = \begin{pmatrix} 1 & \frac{t}{1!} & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \frac{t}{1!} & \dots & \frac{t^{n-2}}{(n-2)!} \\ 0 & 0 & 1 & \dots & \frac{t^{n-3}}{(n-3)!} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix}.$$

So $(B_t)_{ij} = 0$ if $j < i$ and $(B_t)_{ij} = \frac{t^{j-i}}{(j-i)!}$ if $i \le j$.

With $f(\lambda) = e^\lambda$ put $g(\lambda) = f(t\lambda) = e^{t\lambda}$. A simple induction shows that $g^{(j)}(\lambda) = t^j g(\lambda)$. Then

$$g(J_\lambda) = e^{tJ_\lambda} = e^{t\lambda} B_t.$$

In particular,

$$e^{J_\lambda} = e^\lambda B_1.$$

We can even determine the polynomial $p(x)$ for which $p(J_\lambda) = e^{J_\lambda}$. Put

$$p_n(x) = \sum_{j=0}^{n-1} \frac{(x-\lambda)^j}{j!}.$$

It follows easily that

$$p_n^{(j)}(\lambda) = 1 \text{ for } 0 \le j \le n-1, \text{ and } p_n^{(n)}(x) = 0.$$

From Eq. 11.6 we see that

$$p_n(J_n(\lambda)) = B_n(1), \text{ so } e^\lambda p_n(J_n(\lambda)) = e^\lambda B_n(1) = e^{J_\lambda}.$$

We next compute $B_t B_s$, for arbitrary complex numbers t and s. Clearly the product is upper triangular. Then for $i \le j$ we have

$$(B_t \cdot B_s)_{ij} = \sum_{k=i}^{j} (B_t)_{ik} \cdot (B_s)_{kj} = \sum_{k=i}^{j} \frac{t^{k-i}}{(k-i)!} \cdot \frac{s^{j-k}}{(j-k)!}$$

$$= \frac{1}{(j-i)!} \sum_{k=i}^{j} \binom{j-i}{j-k} t^{(j-i)-(j-k)} s^{j-k} = \frac{1}{(j-i)!}(t+s)^{j-i} = (B_{t+s})_{ij}.$$

Hence

$$B_t \cdot B_s = B_{t+s} = B_s \cdot B_t.$$

It now follows that

$$e^{tJ_\lambda} \cdot e^{sJ_\lambda} = e^{t\lambda} B_t \cdot e^{s\lambda} B_s = e^{(t+s)\lambda} B_{t+s} = e^{(t+s)J_\lambda} = e^{sJ_\lambda} \cdot e^{tJ_\lambda}.$$

Fix the complex number s and put $T_0 = \operatorname{diag}(s^{n-1}, s^{n-2}, \ldots, s, 1)$. It is now easy to verify that

$$T_0^{-1} \cdot sJ_\lambda \cdot T_0 = J_{s\lambda}. \tag{11.18}$$

Hence $J_{s\lambda}$ is the Jordan form of sJ_λ. Also, $\left(T_0 A T_0^{-1}\right)_{ij} = s^{j-i}A_{ij}$. From this it follows easily that

$$T_0 B_1 T_0^{-1} = B_s.$$

It now follows that

$$e^{sJ_\lambda} = f(sJ_\lambda) = f(T_0 \cdot J_{s\lambda} \cdot T_0^{-1}) = T_0 \cdot f(J_{s\lambda}) T_0^{-1} =$$
$$= T_0 \cdot e^{J_{s\lambda}} \cdot T_0^{-1} = T_0(e^{s\lambda}B_1)T_0^{-1} = e^{s\lambda}B_s,$$

which agrees with an earlier equation.

Suppose A is a general $n \times n$ matrix with Jordan form

$$T^{-1}AT = J_{\lambda_1} \oplus \cdots \oplus J_{\lambda_t},\ J_{\lambda_i} \in M_{n_i}(F).$$

Let T_0 be the direct sum of the matrices $\mathrm{diag}(s^{n_i-1}, s^{n_i-2}, \ldots, s, 1)$. Then

$$T_0^{-1}T^{-1}sATT_0 = J_{s\lambda_1} \oplus \cdots \oplus J_{s\lambda_t},$$

so

$$e^{sA} = T\left(e^{s\lambda_1}B_{n_1}(s) \oplus \cdots \oplus e^{s\lambda_t}B_{n_t}(s)\right)T^{-1}.$$

Several properties of the exponential function are now easy corollaries.

Corollary 11.5.4. *Using the above descriptions of A and e^{sA}, etc., we obtain:*

(a) $Ae^A = e^A A$ for all square A.

(b) $e^{sA}e^{tA} = e^{(s+t)A} = e^{tA}e^{sA}$, for all $s, t \in \mathcal{C}$.

(c) $e^0 = I$, where 0 is the zero matrix.

(d) e^A is nonsingular and $e^{-A} = (e^A)^{-1}$.

(e) $e^I = eI$.

(f) $det(e^{J_\lambda}) = e^{n\lambda} = e^{trace(J_\lambda)}$, from which it follows that $det(e^A) = e^{trace(A)}$.

11.6 Commuting Matrices*

Let A be a fixed $n \times n$ matrix over $\mathcal{C}$. We want to determine which matrices commute with A. If A commutes with matrices B and C, clearly A commutes with BC and with any linear combination of B and C. Also, if A commutes with every $n \times n$ matrix, then A must be a scalar multiple $A = aI$ of the identity matrix I. (If you have not already verified this, do it now.)

Now suppose

$$T^{-1}AT = J = J_1 \oplus \cdots \oplus J_s \tag{11.19}$$

is the Jordan form of A with each J_i an elementary Jordan block. It is easy to check that X commutes with J if and only if TXT^{-1} commutes with A. Therefore the problem reduces to finding the matrices X that commute with J. Write X in block form corresponding to Eq. 11.19:

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1s} \\ X_{21} & X_{22} & \cdots & X_{2s} \\ \vdots & \vdots & \cdots & \vdots \\ X_{s1} & X_{s2} & \cdots & X_{ss} \end{pmatrix}. \tag{11.20}$$

The condition $JX = XJ$ reduces to the equalities

$$J_p X_{pq} = X_{pq} J_q \quad (p, q = 1, \cdots, s). \tag{11.21}$$

Note that there is only one equation for each X_{pq}. Fix attention on one block $X_{pq} = B = (b_{ij})$. Suppose J_p is $r \times r$ and J_q is $t \times t$. Then $X_{pq} = B$ is $r \times t$, $J_p = \lambda_p I + N_r$, $J_q = \lambda_q I + N_t$. From Eq. 11.21 we easily get

$$\lambda_p B_{uv} + \sum_{k=1}^{r} (N_r)_{uk} B_{kv} \quad = \quad \lambda_q B_{uv} + \sum_{k=1}^{t} B_{uk} (N_t)_{kv}, \tag{11.22}$$

$$\text{for all } p, q, u, v \quad \text{with} \quad 1 \le p, q \le s, 1 \le u \le r;\ 1 \le v \le t.$$

Eq. 11.22 quickly gives the following four equations:

$$(\text{For } v \neq 1,\ u \neq r), \lambda_p B_{uv} + B_{u+1,v} = \lambda_q B_{uv} + B_{u,v-1}. \tag{11.23}$$

$$(\text{For } v = 1; u \neq r), \lambda_p B_{u1} + B_{u+1,1} = \lambda_q B_{u1} + 0. \tag{11.24}$$

$$(\text{For } v \neq 1; u = r), \lambda_p B_{rv} + 0 = \lambda_q B_{r,v} + B_{r,v-1}. \tag{11.25}$$

$$(\text{For } u = r, v = 1), \lambda_p B_{r1} + 0 = \lambda_q B_{r1}. \tag{11.26}$$

First suppose that $\lambda_p \neq \lambda_q$. Then from Eq. 11.26 $B_{r1} = 0$. Then from Eq. 11.24 $(\lambda_q - \lambda_p) B_{u1} = B_{u+1,1}$. Put $u = r-1, r-2, \ldots, 1$, in that order, to

get $B_{u1} = 0$ for $1 \leq u \leq r$, i.e., the first column of B is the zero column. In Eq. 11.25 put $v = 2, 3, \ldots, t$ to get $B_{rv} = 0$ for $1 \leq v \leq t$ i.e., the bottom row of B is the zero row. Put $u = r - 1$ in Eq. 11.23 and then put $v = 2, 3, \ldots, t$, in that order, to get $B_{r-1,v} = 0$ for all v, i.e., the $(r-1)$st column has all entries equal to zero. Now put $u = r-2$ and let $v = 2, 3, \ldots, t$, in that order, to get $B_{r-2,v} = 0$ for $1 \leq v \leq t$. Continue in this way for $u = r - 3,\ r - 4$, etc., to see that $B = 0$.

So for $\lambda_p \neq \lambda_q$ we have $X_{pq} = 0$.

Now consider the case $\lambda_p = \lambda_q$. The equations Eq. 11.23 through 11.25 become

$$(\text{For } v \neq 1; u \neq r), B_{u+1,v} = B_{u,v-1}. \tag{11.27}$$

$$(\text{For } v = 1; u \neq r) B_{u+1,1} = 0. \tag{11.28}$$

$$(\text{For } v \neq 1; u = r) B_{r,v-1} = 0. \tag{11.29}$$

It is now relatively straightforward to check that there are only the following three possibilities (each of which is said to be in *linear triangular form*:

1. $r = t$, in which case $B = \sum_{i=0}^{t} \zeta_i N_t^i$ for some scalars $\zeta_i \in \mathcal{C}$.
2. $r > t$, in which case B has the form

$$B = \begin{pmatrix} \sum_{i=1}^{t} \zeta_i N_t^i \\ 0_{r-t,t} \end{pmatrix}.$$

3. $r < t$, in which case B has the form

$$B = \begin{pmatrix} 0_{r,t-r} & \sum_{i=1}^{r} \zeta_i N_r^i \end{pmatrix}.$$

Conversely, the determination of the form of B shows that if B has any of the forms shown above then B commutes with X. For example, if

$$B = \rho I_3 + N_3 \oplus \rho I_2 + N_2 \oplus \sigma I_3 + N_3 \tag{11.30}$$

for $\rho \neq \sigma$, then the matrices X that commute with V have the form

$$X = \left(\begin{array}{c|c} \sum_{i=1}^{3} a_i N_3^i & \sum_{i=1}^{2} c_i N_2^i \\ & 0_{1,2} \\ 0_{2,1} \quad \sum_{i=1}^{2} d_i N_2^i & \sum_{i=1}^{2} b_i N_2^i \end{array}\right) \oplus \sum_{i=1}^{3} \lambda_i N_3^i,$$

where $a_i, b_i, c_i d_i, \lambda_i$ are arbitrary scalars.

Polynomials in a matrix A have the special property of commuting, not only with the matrix A, but also with any matrix X which commutes with A.

Theorem 11.6.1. *If C commutes with every matrix which commutes with B, then C is a polynomial in B.*

Proof. In the usual way we may assume that B is in Jordan form. Suppose

$$B = \sum_{i=1}^{s} B_i = \sum_{i=1}^{s} J_{n_i}(\lambda_i) = \sum_{i=1}^{s} \lambda_i I_{n_i} + N_{n_i}.$$

The auxiliary matrices

$$X = \sum_{i=1}^{s} a_i I_{n_i}$$

where the a_i are arbitrary numbers, are known to commute with B. So by hypothesis X also commutes with C. From the preceding discussion, if we take the a_i to be distinct, we see that C must be decomposable into blocks:

$$C = C_1 \oplus \cdots \oplus C_s.$$

Moreover, it follows from $CB = BC$ that these blocks have linear triangular form. Now let X be an arbitrary matrix that commutes with B. The general form of the matrix X was established in the preceding section. By assumption, C commutes with X. Representing X in block form, we see that the equality $CX = XC$ is equivalent to the relations

$$C_p X_{pq} = X_{pq} C_q \quad (p, q = 1, 2, \ldots, s). \tag{11.31}$$

If the corresponding blocks B_p, B_q have distinct eigenvalues, then Eq. 11.31 contributes nothing, since then $X_{pq} = 0$. Hence we assume that B_p and B_q have the same eigenvalues. Let

$$C_p = \sum_{i=1}^{r} a_i N_r^i, \; C_q = \sum_{i=1}^{t} b_i N_t^i.$$

For definiteness, suppose that $r < t$. Then Eq. 11.31 becomes

$$\left\{\sum_{i=1}^{r} a_i N_r^i\right\}\cdot\left\{\begin{pmatrix} 0_{r,t-r} & \sum_{i=1}^{r}\zeta_i N_r^i \end{pmatrix}\right\} = \left\{\begin{pmatrix} 0_{r,t-r} & \sum_{i=1}^{r}\zeta_i N_r^i \end{pmatrix}\right\}\cdot\left\{\sum_{i=1}^{t} b_i N_t^i\right\},$$

where the ζ_i are arbitrary complex numbers. Multiply the top row times the right hand column on both sides of the equation to obtain:

$$(a_1,\ldots,a_r)\cdot(\zeta_r,\ldots,\zeta_1)^T = (0,\ldots,0,\zeta_1,\ldots,\zeta_r)\cdot(b_t,\ldots,b_1)^T,$$

which is equivalent to

$$a_1\zeta_r + a_2\zeta_{r-1} + \cdots + a_r\zeta_1 = b_1\zeta_r + b_2\zeta_{r-1} + \cdots + b_r\zeta_1.$$

Since the ζ_i can be chosen arbitrarily, it must be that

$$a_i = b_i, \quad \text{for } 1 \le i \le r. \tag{11.32}$$

It is routine to verify that a similar equality is obtained for $r \ge t$.

Suppose that the Jordan blocks of the matrix B are such that blocks with the same eigenvalues are adjacent. For example, let

$$B = (B_1 \oplus \cdots \oplus B_{m_1}) \oplus (B_{m_1+1} \oplus \cdots \oplus B_{m_2}) \oplus \cdots \oplus (B_{m_k+1} \oplus \cdots \oplus B_s),$$

where blocks with the sam eigenvalues are contained in the same set of parentheses. Denoting the sums in parentheses by $B^{(1)}, B^{(2)}, \cdots$, respectively , divide the matrix B into larger blocks which we call *cells*. Then divide the matrices X and C into cells in a corresponding way. The results above show that all the nondiagonal cells of the matrix X are zero; the structure of the diagonal cells was also described above. The conditions for the matrix C obtained in the present section show that the diagonal blocks of this matrix decompose into blocks in the same way as those of B. Moreover, the blocks of the matrix C have a special triangular form. Equalities 11.32 mean that in the blocks of the matrix C belonging to a given cell, the nonzero elements lying on a line parallel to the main diagonal are equal.

For example, let B have the form given in Eq. 11.30,

$$B = \begin{pmatrix} \rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma \end{pmatrix}. \tag{11.33}$$

Our results show that every matrix C that commutes with each matrix which commutes with B must have the form

$$C = \begin{pmatrix} a_0 & a_1 & a_2 & & & & & \\ & a_0 & a_1 & & & & & \\ & & a_0 & & & & & \\ & & & a_0 & a_1 & & & \\ & & & & a_0 & & & \\ & & & & & d_0 & d_1 & d_2 \\ & & & & & & d_0 & d_1 \\ & & & & & & & d_0 \end{pmatrix}. \tag{11.34}$$

We need to prove that C can be represented as a polynomial in B. We do this only for the particular case where B has the special form given in Eq. 11.33, so that C has its form given in Eq. 11.34. By Lemma 11.4.1 there is a polynomial $f(\lambda)$ that satisfies the conditions:

$$\begin{aligned} f(\rho) = a_0, \quad f'(\rho) = a_1, \quad f''(\rho) = a_2, \\ f(\sigma) = d_0, \quad f'(\sigma) = d_1, \quad f''(\sigma) = d_2. \end{aligned}$$

By applying Eq. 11.6 we see that $f(B) = C$ □

11.7 A Matrix Differential Equation*

Let F denote either $\mathcal{R}$ or $\mathcal{C}$. If $B(t)$ is an $n \times n$ matrix each of whose entries is a differentiable function $b_{ij}(t)$ from F to F, we say that *the derivative of B with respect to t* is the matrix $\frac{d}{dt}(B(t))$ whose (i,j)th entry is $(b_{ij})'(t)$.

Let $x(t) = (x_1(t), \ldots, x_n(t))^T$ be an n-tuple of unknown functions $F \to F$. The derivative of $x(t)$ will be denoted $\dot{x}(t)$. Let A be an $n \times n$ matrix over F, and consider the matrix differential equation (with initial condition):

$$\dot{x}(t) = Ax(t),\ x(0) = x_0. \tag{11.35}$$

Theorem 11.7.1. *The solution to the system of differential equations in Eq. 11.35 is given by* $x(t) = e^{tA}x_0$.

Before we can prove this theorem, we need the following lemma.

Lemma 11.7.2.

$$\frac{d}{dt}(e^{tA}) = Ae^{tA}.$$

Proof. First suppose that A is the elementary Jordan block $A = J = J_n(\lambda) = \lambda I + N$, where $N = N_n$. If $n = 1$, recall that the ordinary differential equation $x'(t) = \lambda x(t)$ has the solution $x(t) = ae^{\lambda t}$ where $a = x(0)$. Since $N_1 = 0$, we see that the theorem holds in this case. Now suppose that $n > 1$. So $A = J = \lambda I + N$ where $N^n = 0$. Recall that

$$e^A = e^{tJ} = e^{\lambda t}B_t = \sum_{j=0}^{n-1} \frac{e^{\lambda t} \cdot t^j}{j!} N^j.$$

A simple calculation shows that

$$\frac{d}{dt}\left(e^{tJ}\right) = \lambda e^{\lambda t} I + \sum_{j=1}^{n-1} \left(\frac{\lambda e^{\lambda t} t^j + j e^{\lambda t} t^{j-1}}{j!} \right) N^j.$$

On the other hand,

$$Je^{tJ} = (\lambda I + N)\left(e^{\lambda t} \sum_{j=0}^{n-1} \frac{t^j}{j!} N^j \right)$$

$$= e^{\lambda t}\left[(\lambda I + N)\left(I + \sum_{j=1}^{n-2} \frac{t^j}{j!} N^j + \frac{t^{n-1}}{j!} N^{n-1} \right) \right]$$

$$= e^{\lambda t}\left[\lambda I + (1 + \lambda t)N + \sum_{j=2}^{n-2} \left(\frac{\lambda t^j}{j!} + \frac{t^{j-1}}{(j-1)!} \right) N^j + \right.$$

$$\left. + \left(\frac{\lambda t^{n-1}}{(n-1)!} + \frac{t^{n-2}}{(n-2)!} \right) N^{n-1}) \right]$$

$$= \frac{d}{dt}\left(e^{tJ}\right), \text{ as desired.}$$

This completes a proof of the Lemma. Now turn to a proof of Theorem 11.7.1.

First suppose that A has the Jordan form $J = P^{-1}AP$, where $J = J_1 \oplus \cdots \oplus J_s$. Since e^{tA} is a polynomial in tA, $e^{P^{-1}tAP} = P^{-1}e^{tA}P$, i.e.

$$e^{tA} = P \cdot e^{tJ} \cdot P^{-1} = T \cdot (e^{tJ_1} \oplus \cdots \oplus e^{tJ_s})P^{-1}.$$

It now follows easily that

$$\frac{d}{dt}\left(e^{tA}\right) = Ae^{tA}.$$

Since $x_0 = x(0)$ is a constant matrix, $\frac{d}{dt}\left(e^{tA}x_0\right) = A^{tA}x_0$, so that $x = e^{tA}x_0$ satisfies the differential equation $\dot{x} = Ax$. □

It is a standard result from the theory of differential equations that this solution is the unique solution to the given equation.

11.8 Exercises

1. Show that the norm of $T \in \mathcal{L}(V, W)$ as defined in Equation 11.1 is a vector norm on $\mathcal{L}(V, W)$ viewed as a vector space in the usual way.

2. Let D denote the derivative operator, and for a function f (in our case a formal power series or Laurent series) let $f^{(j)}$ denote the jth derivative of f, i.e., $D^j(f) = f^{(j)}$.

   (i) Prove that
   $$D^n(f \cdot g) = \sum_{i=0}^{n} \binom{n}{i} f^{(i)} g^{(n-i)}.$$

   (ii) Derive as a corollary to part (i) the fact that
   $$D^j(f^2) = \sum_{i_1+i_2=j} \binom{j}{i_1, i_2} f^{(i_1)} f^{(i_2)}.$$

   (iii) Now use part (i) and induction on n to prove that
   $$D^j(f^n) = \sum_{i_1+\cdots+i_n=j} \binom{j}{i_1, \ldots, i_n} f^{(i_1)} \cdots f^{(i_n)}.$$

3. The *Frobenius norm* $\|A\|_F$ of an $m \times n$ matrix A is defined to be the square root of the sum of the squares of the magnitudes of all the entries of A.

   (a) Show that the Frobenius norm really is a matrix norm (i.e., a vector norm on the vector space of all $m \times n$ matrices).

   (b) Compute $\|I\|_F$ and deduce that $\|\cdot\|_F$ cannot be a transformation norm induced by some vector norm.

   (c) Let U and V be unitary matrices of the appropriate sizes and show that
   $$\|A\|_F = \|UA\|_F = \|AV\|_F = \|UAV\|_F.$$

4. Let $\lambda \mapsto f(\lambda)$ and $\lambda \mapsto g(\lambda)$ be two numerical functions and let A be an $n \times n$ matrix for which both $f(A)$ and $g(A)$ are defined. Show that $f(A)g(A) = g(A)f(A)$.

Chapter 12

Infinite Dimensional Vector Spaces*

12.1 Partially Ordered Sets & Zorn's Lemma*

There are occasions when we would like to indulge in a kind of "infinite induction." Basically this means that we want to show the existence of some set which is maximal with respect to certain specified properties. In this text we want to use Zorn's Lemma to show that every vector space has a basis, i.e., a maximal linearly independent set of vectors in some vector space. The maximality is needed to show that these vectors span the entire space.

In order to state Zorn's Lemma we need to set the stage.

Definition A *partial order* on a nonempty set Z is a relation $\leq$ on A satisfying the following:

1. $x \leq x$ for all $x \in A$ (reflexive) ;
2. if $x \leq y$ and $y \leq x$ then $x = y$ for all $x, y \in A$ (antisymmetric);
3. if $x \leq y$ and $y \leq z$ then $x \leq z$ for all $x, y, z \in A$ (transitive).

Given a partial order $\leq$ on A we often say that A is partially ordered by $\leq$, or that $(A, \leq)$ is a *partially ordered set.*

Definition Let the nonempty set A be partially ordered by $\leq$.

1. A subset V of A is called a *chain* if for all $x, y \in B$, either $x \leq y$ or $y \leq x$.

2. An *upper bound* for a subset B of A is an element $u \in A$ such that $b \leq u$ for all $b \in B$.

3. A *maximal element* of A is an element $m \in A$ such that if $m \leq x$ for some $x \in A$, then $m = x$.

In the literature there are several names for chains, such as *linearly ordered subset* or *simply ordered subset*. The existence of upper bounds and maximal elements depends on the nature of $(A, \leq)$.

As an example, let A be the collection of all *proper* subsets of $\mathcal{Z}^+$ (the set of positive integers) ordered by $\subseteq$. Then, for example, the chain

$$\{1\} \subseteq \{1, 2\} \subseteq \{1, 2, 3\} \subseteq \cdots$$

does not have an upper bound. However, the set A does have maximal elements: for example $\mathcal{Z}^+ \setminus \{n\}$ is a maximal element of A for any $n \in \mathcal{Z}^+$.

Zorn's Lemma If A is a nonempty partially ordered set in which every chain has an upper bound, then A has a maximal element. It is a nontrivial result that *Zorn's Lemma is independent of the usual (Zermelo-Fraenkel) axioms of set theory* in the sense that if the axioms of set theory are consistent, then so are these axioms together with Zorn's Lemma or with the negation of Zorn's Lemma. The two other most nearly standard axioms that are equivalent to Zorn's Lemma (in the presence of the usual Z-F axioms) are the *Axiom of Choice* and the *Well Ordering Principle*. In this text, we just use Zorn's Lemma and leave any further discussion of these matters to others.

12.2 Bases for Vector Spaces*

Let V be any vector space over an arbitrary field. Let $S = \{A \subseteq V : A \text{ is linearly independent}\}$ and let S be partially ordered by inclusion. If C is any chain in S, then the union of all the sets in C is an upper bound for C. Hence by Zorn's Lemma S must have a maximal element $\mathcal{B}$. By definition $\mathcal{B}$ is a linearly independent set not properly contained in any other linearly independent set. If a vector v were not in the space spanned by $\mathcal{B}$, then $\mathcal{B} \cup \{v\}$ would be a linearly independent set properly containing $\mathcal{B}$, a contradiction. Hence $\mathcal{B}$ is a basis for V. This proves the following theorem:

Theorem 12.2.1. *Each vector space V over an arbitrary field F has a basis. This is a subset $\mathcal{B}$ of V such that each vector $v \in V$ can be written in just one way as a linear combination of a finite lis of vectors in $\mathcal{B}$.*

12.3 A Theorem of Philip Hall*

Let $\mathcal{S}$ and I be arbitrary sets. For each $i \in I$ let $A_i \subseteq \mathcal{S}$. If $a_i \in A_i$ for all $i \in I$, we say $\{a_i : i \in I\}$ is a *system of representatives* for $\mathcal{A} = (A_i : i \in I)$. If in addition $a_i \neq a_j$ whenever $i \neq j$, even though A_i may equal A_j, then $\{a_i : i \in I\}$ is a *system of distinct representatives* (SDR) for $\mathcal{A}$. Our first problem is: Under what conditions does some family $\mathcal{A}$ of subsets of a set $\mathcal{S}$ have an SDR?

For a finite collection of sets a reasonable answer was given by Philip Hall in 1935. It is obvious that if $\mathcal{A} = (A_i : i \in I)$ has an SDR, then the union of each k of the members of $\mathcal{A} = (A_i : i \in I)$ must have at least k elements. Hall's observation was that this obvious necessary condition is also sufficient. We state the condition formally as follows:

Condition (H) : Let $I = [n] = \{1, 2, \ldots, n\}$, and let S be any (nonempty) set. For each $i \in I$, let $A_i \subseteq S$. Then $\mathcal{A} = (S_1, \ldots, S_n)$ satisfies **Condition (H)** provided for each $K \subseteq I$, $|\cup_{k\in K} S_k| \geq |K|$.

Theorem 12.3.1. *The family $\mathcal{A} = (S_1, \ldots, S_n)$ of finitely many (not necessarily distinct) sets has an SDR if and only if it satisfies Condition (H).*

Proof. As Condition (H) is clearly necessary, we now show that it is also sufficient. $B_{r,s}$ denotes a block of r subsets $(S_{i_1}, \ldots, S_{i_r})$ belonging to $\mathcal{A}$, where $s = |\cup\{S_j : S_j \in B_{r,s}\}|$. So Condition (H) says: $s \geq r$ for each block $B_{r,s}$. If $s = r$, $B_{r,s}$ is called a *critical block.* (By convention, the empty block $B_{0,0}$ is *critical.*)

If $B_{r,s} = (A_1, \ldots, A_u, C_{u+1}, \ldots, C_r)$ and
$B_{t,v} = (A_1, \ldots, A_u, D_{u+1}, \ldots, D_t)$, write $B_{r,s} \cap B_{t,v} =$
$(A_1, \ldots, A_u)$; $B_{r,s} \cup B_{t,v} = (A_1, \ldots, A_u, C_{u+1}, \ldots, C_r, D_{u+1}, \ldots, D_t)$. Here the notation implies that $A_1, \ldots, A_u$ are precisely the subsets in both blocks. Then write

$B_{r,s} \cap B_{t,v} = B_{u,w}$, where $w = |\cup\{A_i : 1 \leq i \leq u\}|$, and $B_{r,s} \cup B_{t,v} = B_{y,z}$, where $y = r + t - u$, $z = |\cup\{S_i : S_i \in B_{r,s} \cup B_{t,v}\}|$.

The proof will be by induction on the number n of sets in the family $\mathcal{A}$, but first we need two lemmas.

Lemma 12.3.2. *If $\mathcal{A}$ satisfies Condition (H), then the union and intersection of critical blocks are themselves critical blocks.*

Proof of Lemma 12.3.2. Let $B_{r,r}$ and $B_{t,t}$ be given critical blocks. Say $B_{r,r} \cap B_{t,t} = B_{u,v}$; $B_{r,r} \cup B_{t,t} = B_{y,z}$. The z elements of the union will be the $r + t$ elements of $B_{r,r}$ and $B_{t,t}$ reduced by the number of elements in both blocks, and this latter number includes at least the v elements in the intersection: $z \leq r + t - v$. Also $v \geq u$ and $z \geq y$ by Condition (H). Note: $y + u = r + t$. Hence $r + t - v \geq z \geq y = r + t - u \geq r + t - v$, implying that equality holds throughout. Hence $u = v$ and $y = z$ as desired for the proof of Lemma 12.3.2 .

Lemma 12.3.3. *If $B_{k,k}$ is any critical block of $\mathcal{A}$, the deletion of elements of $B_{k,k}$ from all sets in $\mathcal{A}$ not belonging to $B_{k,k}$ produces a new family $\mathcal{A}'$ in which Condition (H) is still valid.*

Proof of Lemma12.3.3. Let $B_{r,s}$ be an arbitrary block, and $(B_{r,s})' = B'_{r,s'}$ the block after the deletion. We must show that $s' \geq r$. Let $B_{r,s} \cap B_{k,k} = B_{u,v}$ and $B_{r,s} \cup B_{k,k} = B_{y,z}$. Say

$$B_{r,s} = (A_1, \ldots, A_u, C_{u+1}, \ldots, C_r),$$

$$B_{k,k} = (A_1, \ldots, A_u, D_{u+1}, \ldots, D_k).$$

So $B_{u,v} = (A_1, \ldots, A_u)$, $B_{y,z} = (A_1, \ldots, A_u, C_{u+1}, \ldots, C_r, D_{u+1}, \ldots, D_k)$. The deleted block $(B_{r,s})' = B'_{r,s'}$ is $(A_1, \ldots, A_u, C'_{u+1}, \ldots, C'_r)$. But $C_{u+1}, \ldots, C_r$, as blocks of the union $B_{y,z}$, contain at least $z - k$ elements not in $B_{k,k}$. Thus $s' \geq v + (z - k) \geq u + y - k = u + (r + k - u) - k = r$. Hence $s' \geq r$, as desired for the proof of Lemma 12.3.3.

As indicated above, for the proof of the main theorem we now use induction on n. For $n = 1$ the theorem is obviously true.

Induction Hypothesis: Suppose the theorem holds (Condition (H) implies that there is an SDR) for any family of m sets, $1 \leq m < n$.

We need to show the theorem holds for a system of n sets. So let $1 < n$, assume the induction hypothesis, and let $\mathcal{A} = (S_1, \ldots, S_n)$ be a given collection of subsets of S satisfying Condition (H).

First Case: There is some critical block $B_{k,k}$ with $1 \leq k < n$. Delete the elements in the members of $B_{k,k}$ from the remaining subsets, to obtain a new family $\mathcal{A}' = B_{k,k} \cup B'_{n-k,v}$, where $B_{k,k}$ and $B'_{n-k,v}$ have no common elements in their members. By Lemma 12.3.3, Condition (H) holds in $\mathcal{A}'$, and hence holds separately in $B_{k,k}$ and in $B'_{n-k,v}$ viewed as families of sets.

By the induction hypothesis, $B_{k,k}$ and $B'_{n-k,v}$ have (disjoint) SDR's whose union is an SDR for $\mathcal{A}$.

Remaining Case: There is no critical block for $\mathcal{A}$ except possibly the entire system. Select any S_j of $\mathcal{A}$ and then select any element of S_j as its representative. Delete this element from all remaining sets to obtain a family $\mathcal{A}'$. Hence a block $B_{r,s}$ with $r < n$ becomes a block $B'_{r,s'}$ with $s' \in \{s, s-1\}$. By hypothesis $B_{r,s}$ was not critical, so $s \geq r+1$ and $s' \geq r$. So Condition (H) holds for the family $\mathcal{A}' \setminus \{S_j\}$, which by induction has an SDR. Add to this SDR the element selected as a representative for S_j to obtain an SDR for $\mathcal{A}$. □

We now interpret the SDR problem as one on matchings in bipartite graphs. Let $G = (X,\ Y,\ E)$ be a bipartite graph. For each $S \subseteq X$, let $N(S)$ denote the set of elements of Y connected to at least one element of S by an edge, and put $\delta(S) = |S| - |N(S)|$. Put $\delta(G) = max\{\delta(S) : S \subseteq X\}$. Since $\delta(\emptyset) = 0$, clearly $\delta(G) \geq 0$. Then Hall's theorem states that G has an X-saturating matching if and only if $\delta(G) = 0$.

Theorem 12.3.4. *G has a matching of size t (or larger) if and only if $t \leq |X| - \delta(S)$ for all $S \subseteq X$.*

Proof. First note that Hall's theorem says that G has a matching of size $t = |X|$ if and only if $\delta(S) \leq 0$ for all $S \subseteq X$ iff $|X| \leq |X| - \delta(S)$ for all $S \subseteq X$. So our theorem is true in case $t = |X|$. Now suppose that $t < |X|$. Form a new graph $G' = (X,\ Y \cup Z,\ E')$ by adding new vertices $Z = \{z_1, \ldots, z_{|X|-t}\}$ to Y, and join each z_i to each element of X by an edge of G'.

If G has a matching of size t, then G' has a matching of size $|X|$, implying that for all $S \subseteq X$,

$$|S| \leq |N'(S)| = |N(S)| + |X| - t,$$

implying

$$|N(S)| \geq |S| - |X| + t = t - (|X| - |S|) = t - |X \setminus S|.$$

This is also equivalent to $t \leq |X| - (|S| - |N(S)|) = |X| - \delta(S)$.

Conversely, suppose $|N(S)| \geq t - |X \setminus S| = t - (|X| - |S|)$. Then $|N'(S)| = |N(S)| + |X| - t \geq (t - |X| + |S|) + |X| - t = |S|$. By Hall's theorem, G' has an X-saturating matching M. At *most* $|X| - t$ edges of M join X to Z, so at *least* t edges of M are from X to Y. □

Note that $t \leq |X| - \delta(S)$ for all $S \subseteq X$ iff $t \leq min_{S \subseteq X}(|X| - \delta(S)) = |X| - max_{S \subseteq X}\delta(S) = |X| - \delta(G)$.

Corollary 12.3.5. *The largest matching of G has size $|X| - \delta(G) = m(G)$, i.e., $m(G) + \delta(G) = |X|$.*

12.4 A Theorem of Marshall Hall, Jr.*

Many of the ideas of "finite" combinatorics have generalizations to situations in which some of the sets involved are infinite. We just touch on this subject.

Given a family $\mathcal{A}$ of sets, if the number of sets in the family is infinite, there are several ways the theorem of P. Hall can be generalized. One of the first (and to our mind one of the most useful) was given by Marshal Hall, Jr. (no relative of P. Hall), and is as follows.

Theorem 12.4.1. *Suppose that for each i in some index set I there is a* finite *subset A_i of a set S. The system $\mathcal{A} = (A_i)_{i \in I}$ has an SDR if and only if the following Condition (H') holds: For each finite subset I' of I the system $\mathcal{A}' = (A_i)_{i \in I'}$ satisfies Condition (H).*

Proof. We establish a partial order on deletions, writing $D_1 \subseteq D_2$ for deletions D_1 and D_2 iff each element deleted by D_1 is also deleted by D_2. Of course, we are interested only in deletions which preserve Condition (H'). If all deletions in an ascending chain $D_1 \subseteq D_2 \subseteq \cdots \subseteq D_i \subseteq \cdots$ preserve Condition (H), let D be the deletion which consists of deleting an element b from a set A iff there is some i for which b is deleted from A by D_i. *We assert that deletion D also preserves Condition (H).*

In any block $B_{r,s}$ of $\mathcal{A}$, $(r, s < \infty)$, at most a finite number of deletions in the chain can affect $B_{r,s}$. If no deletion of the chain affects $B_{r,s}$, then of course D does not affect $B_{r,s}$, and Condition (H) still holds for $B_{r,s}$. Otherwise, let D_n be the last deletion that affects $B_{r,s}$. So under D_n (and hence also under D) $(B_{r,s})' = B'_{r,s'}$ still satisfies Condition (H) by hypothesis, i.e., $s' \geq r$. But $B_{r,s}$ is arbitrary, so D preserves Condition (H) on $\mathcal{A}$. By Zorn's Lemma,

there will be a maximal deletion $\bar{D}$ preserving Condition (H). We show that under such a maximal deletion $\bar{D}$ preserving Condition H, each deleted set S_i' has only a single element. Clearly these elements would form an SDR for the original $\mathcal{A}$.

Suppose there is an a_1 not belonging to a critical block. Delete a_1 from every set A_i containing a_1. Under this deletion a block $B_{r,s}$ is replaced by a block $B'_{r,s'}$ with $s' \geq s-1 \geq r$, so Condition (H) is preserved. Hence after a maximal deletion each element left is in some critical block. And if $B_{k,k}$ is a critical block, we may delete elements of $B_{k,k}$ from all sets not in $B_{k,k}$ and still preserve Condition (H) by Lemma 12.3.3 (since it needs to apply only to finitely many sets at a time). By Theorem 12.3.1 each critical block $B_{k,k}$ (being finite) possesses an SDR when Condition (H) holds. Hence we may perform an additional deletion leaving $B_{k,k}$ as a collection of singleton sets and with Condition (H) still holding for the entire remaining sets. It is now clear that after a maximal deletion $\bar{D}$ preserving Condition (H), each element is in a critical block, and each critical block consists of singleton sets. Hence after a maximal deletion $\bar{D}$ preserving Condition (H), each set consists of a single element, and these elements form an SDR for $\mathcal{A}$. □

The following theorem, sometimes called the Cantor–Schroeder–Bernstein Theorem, will be used with the theorem of M. Hall, Jr. to show that any two bases of a vector space V over a field F must have the same cardinality.

Theorem 12.4.2. *Let X, Y be sets, and let $\theta : X \to Y$ and $\psi : Y \to X$ be injective mappings. Then there exists a bijection $\phi : X \to Y$.*

Proof. The elements of X will be referred to as males, those of Y as females. For $x \in X$, if $\theta(x) = y$, we say y is the daughter of x and x is the father of y. Analogously, if $\psi(y) = x$, we say x is the son of y and y is the mother of x. A male with no mother is said to be an "adam." A female with no father is said to be an "eve." Ancestors and descendants are defined in the natural way, except that each x or y is both an ancestor of itself and a descendant of itself. If $z \in X \cup Y$ has an ancestor that is an adam (resp., eve) we say that z has an adam (resp., eve). Partition X and Y into the following disjoint sets:

$$X_1 = \{x \in X : x \text{ has no } \textit{eve}\};$$

$$X_2 = \{x \in X : x \text{ has an } \textit{eve}\};$$

$$Y_1 = \{y \in Y : y \text{ has no } eve\};$$

$$Y_2 = \{y \in Y : y \text{ has an } eve\}.$$

Now a little thought shows that $\theta : X_1 \to Y_1$ is a bijection, and $\psi^{-1} : X_2 \to Y_2$ is a a bijection. So

$$\phi = \theta|_{X_1} \cup \psi^{-1}|_{X_2}$$

is a bijection from X to Y. □

Corollary 12.4.3. *If V is a vector space over the field F and if B_1 and B_2 are two bases for V, then $|B_1| = |B_2|$.*

Proof. Let $B_1 = \{x_i : i \in I\}$ and $B_2 = \{y_j : j \in J\}$. For each $i \in I$, let $\Gamma_i = \{j \in J : y_j$ occurs with nonzero coefficient in the unique linear expression for x_i in terms of the $y_j's\}$. Then the union of any k (≥ 1) Γ_i's, say $\Gamma_{i_1}, \ldots, \Gamma_{i_k}$, each of which of course is finite, must contain at least k distinct elements. For otherwise $x_{i_1}, \ldots, x_{i_k}$ would belong to a space of dimension less than k, and hence be linearly dependent. Thus the family $(\Gamma_i : i \in I)$ of sets must have an SDR. This means there is a function $\theta : I \to J$ which is an injection. Similarly, there is an injection $\psi : J \to I$. So by the preceding theorem there is a bijection $J \leftrightarrow I$, i.e., $|B_1| = |B_2|$. □

12.5 Exercises*

Exercise 12.5.0.1. *Let $\mathcal{A} = (A_1, \ldots, A_n)$ be a family of subsets of $\{1, \ldots, n\}$. Suppose that the incidence matrix of the family is invertible. Show that the family has an SDR.*

Exercise 12.5.0.2. *Prove the following generalization of Hall's Theorem:*

Let $\mathcal{A} = (A_1, \ldots, A_n)$ be a family of subsets of X that satisfies the following property: There is an integer r with $0 \leq r < n$ for which the union of each subfamily of k subsets of $\mathcal{A}$, for all k with $0 \leq k \leq n$, has at least $k - r$ elements. Then there is a subfamily of size $n - r$ which has an SDR. (Hint: Start by adding r "dummy" elements that belong to all the sets.)

Exercise 12.5.0.3. *Let G be a (finite, undirected, simple) graph with vertex set V. Let $C = \{C_x : x \in V\}$ be a family of sets indexed by the vertices of G. For $X \subseteq V$, let $C_X = \cup_{x \in X} C_x$. A set $X \subseteq V$ is C-colorable if one can assign to each vertex $x \in X$ a "color" $c_x \in C_x$ so that $c_x \neq c_y$ whenever x and y are adjacent in G. Prove that if $|C_X| \geq |X|$ whenever X induces a connected subgraph of G, then V is C-colorable. (In the current literature of graph theory, the sets assigned to the vertices are called* lists*, and the desired proper coloring of G chosen from the lists is a* list coloring *of G. When G is a complete graph, this exercise gives precisely Hall's Theorem on SDR's. A current research topic in graph theory is the investigation of modifications of this condition that suffice for the existence of list colorings.*

Exercise 12.5.0.4. *With the same notation of the previous exercise, prove that if every proper subset of V is C-colorable and $|C_V| \geq |V|$, then V is C-colorable.*

Index

$A = [T]_{\mathcal{B}_2,\mathcal{B}_1}$, 33
QR-decomposition, 166
$[v]_{\mathcal{B}}$, 32
∞-norm, 123
$\mathcal{L}(U,V)$, 28
n-linear, 55
1-norm, 123
2-norm, 123

algebraic multiplicity of an eigenvalue, 179
algebraically closed, 50
alternating (n-linear), 58

basis of a vector space, 22
basis of quotient space, 184
bijective, 39
binomial theorem, 51

Cayley-Hamilton Theorem, 75
chain, 225
characteristic polynomial, 71
classical adjoint, 70
companion matrix, 73
complex number, 8
convergence of a sequence of matrices, 212
convergence of a sequence of vectors, 123
convergence of power series in a matrix, 211
coordinate matrix, 32

determinant, 59
diagonalizable, 99
dimension, 23
direct sum of subspaces, 16
dual basis, 38
dual space, 38, 130

eigenvalue, 95
eigenvector, 96
elementary divisors of a matrix, 187
elementary Jordan block, 187
exponential of a matrix, 212, 216

finite dimensional, 22
Frobenius norm of a matrix, 224
Fundamental Theorem of Algebra, 50

generalized eigenvector, 178
geometric multiplicity of an eigenvalue, 163
Gram-Schmidt algorithm, 124

Hermitian operator, 139
hyperplane, 171

ideal in $F[x]$, 49
idempotent, 31
independent sets of subspaces, 16
injective, 29

inner product, 115
inner product space, 116
isometry, 151
isomorphic, 39, 46

Jordan form, 186

kernel, 29
Kronecker delta, 8

Lagrange Interpolation, 46
Lagrange interpolation generalized, 208
Laplace expansion, 61, 69
least-squares solution, 168
linear algebra, 41
linear functional, 38, 130
linear operator, 33
linear transformation, 27
linear triangular form, 218
linearly dependent, 20
linearly independent, 20
list, 19
lnull(A), 30

maximal element, 226
monic polynomial, 43
Moore-Penrose generalized inverse, 164

nilpotent, 178
norm, 119
normal operator, 143
normal matrix, 143
normalized QR-decomposition, 166
null space, 29
nullity, 30

orthogonal, 117
orthogonal matrix, 126
orthonormal, 117

partial order, 225
partially ordered set, 225
polarization identity, complex, 136
polarization identity, real, 136
polynomial, 43
positive (semidefinite) operator, 149
prime polynomial, 52
principal ideal, 49
projection, 31
projection matrix, 166
pseudoinverse, 164

Rayleigh Principle, 133
reduced singular value decomposition, 167
rnull(A), 30

scalar matrix functions, 206
Schur's Theorem, 125
self-adjoint operator, 139
singular value decomposition, 158
skew-symmetric matrix, 143
span, 20
sum of subspaces, 16
surjective, 30
system of distinct representatives (SDR), 227

Taylor's Formula, 52
tensor product, 78
trace of a matrix, 72
transformation norm, 199

unitary matrix, 126
upper bound, 226

vector space, 13
vector subspace, 15

Zorn's Lemma, 226