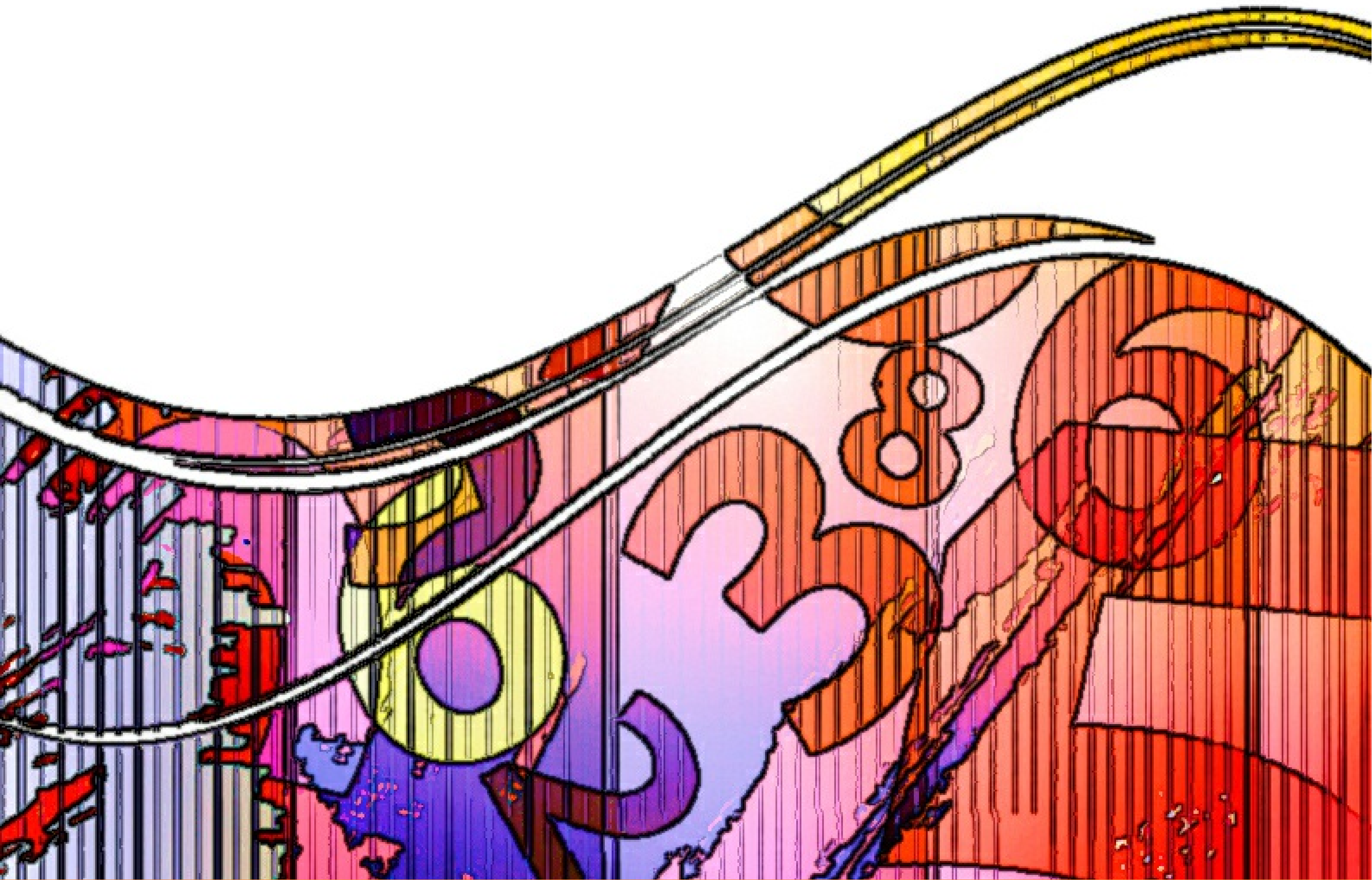
Kenneth Kuttler

Linear Algebra II

Spectral Theory and Abstract Vector Spaces



bookboon
The eBook company

KENNETH KUTTLER

LINEAR ALGEBRA II

SPECTRAL THEORY AND ABSTRACT VECTOR SPACES

Linear Algebra II: Spectral Theory and Abstract Vector Spaces
2nd edition

ISBN 978-87-403-3164-6

CONTENTS

Chapter 7

Vector Spaces And Fields

7.1 Vector Space Axioms

It is time to consider the idea of a Vector space.

Definition 7.1.1 *A vector space is an Abelian group of "vectors" satisfying the axioms of an Abelian group,*

$$\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v},$$

the commutative law of addition,

$$(\mathbf{v}+\mathbf{w})+\mathbf{z}=\mathbf{v}+(\mathbf{w}+\mathbf{z}),$$

the associative law for addition,

$$\mathbf{v}+\mathbf{0}=\mathbf{v},$$

the existence of an additive identity,

$$\mathbf{v}+(-\mathbf{v})=\mathbf{0},$$

the existence of an additive inverse, along with a field of "scalars", $\mathbb{F}$ which are allowed to multiply the vectors according to the following rules. (The Greek letters denote scalars.)

$$\alpha(\mathbf{v}+\mathbf{w})=\alpha\mathbf{v}+\alpha\mathbf{w}, \tag{7.1}$$

$$(\alpha+\beta)\mathbf{v}=\alpha\mathbf{v}+\beta\mathbf{v}, \tag{7.2}$$

$$\alpha(\beta\mathbf{v})=\alpha\beta(\mathbf{v}), \tag{7.3}$$

$$1\mathbf{v}=\mathbf{v}. \tag{7.4}$$

The field of scalars is usually $\mathbb{R}$ or $\mathbb{C}$ and the vector space will be called real or complex depending on whether the field is $\mathbb{R}$ or $\mathbb{C}$. However, other fields are also possible. For example, one could use the field of rational numbers or even the field of the integers mod p for p a prime. A vector space is also called a linear space.

For example, $\mathbb{R}^n$ with the usual conventions is an example of a real vector space and $\mathbb{C}^n$ is an example of a complex vector space. Up to now, the discussion has been for $\mathbb{R}^n$ or $\mathbb{C}^n$ and all that is taking place is an increase in generality and abstraction.

There are many examples of vector spaces.

Example 7.1.2 *Let Ω be a nonempty set and let V consist of all functions defined on Ω which have values in some field $\mathbb{F}$. The vector operations are defined as follows.*

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ (\alpha f)(x) &= \alpha f(x)\end{aligned}$$

Then it is routine to verify that V with these operations is a vector space.

Note that $\mathbb{F}^n$ actually fits in to this framework. You consider the set Ω to be $\{1,2,\cdots,n\}$ and then the mappings from Ω to $\mathbb{F}$ give the elements of $\mathbb{F}^n$. Thus a typical vector can be considered as a function.

Example 7.1.3 *Generalize the above example by letting V denote all functions defined on Ω which have values in a vector space W which has field of scalars $\mathbb{F}$. The definitions of scalar multiplication and vector addition are identical to those of the above example.*

7.2 Subspaces And Bases

7.2.1 Basic Definitions

Definition 7.2.1 *If $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \subseteq V$, a vector space, then*

$$\text{span}\left(\mathbf{v}_1, \cdots, \mathbf{v}_n\right) \equiv \left\{\sum_{i=1}^{n} \alpha_i \mathbf{v}_i : \alpha_i \in \mathbb{F}\right\}.$$

A subset, $W \subseteq V$ is said to be a subspace if it is also a vector space with the same field of scalars. Thus $W \subseteq V$ for W nonempty is a subspace if $ax + by \in W$ whenever $a, b \in \mathbb{F}$ and $x, y \in W$. The span of a set of vectors as just described is an example of a subspace.

Example 7.2.2 *Consider the real valued functions defined on an interval $[a, b]$. A subspace is the set of continuous real valued functions defined on the interval. Another subspace is the set of polynomials of degree no more than 4.*

Definition 7.2.3 *If $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \subseteq V$, the set of vectors is linearly independent if*

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

implies

$$\alpha_1 = \cdots = \alpha_n = 0$$

and $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is called a basis for V if

$$\text{span}\left(\mathbf{v}_1, \cdots, \mathbf{v}_n\right) = V$$

and $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is linearly independent. The set of vectors is linearly dependent if it is not linearly independent.

7.2.2 A Fundamental Theorem

The next theorem is called the exchange theorem. It is very important that you understand this theorem. It is so important that I have given several proofs of it. Some amount to the same thing, just worded differently.

Theorem 7.2.4 *Let $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ be a linearly independent set of vectors such that each $\mathbf{x}_i$ is in the span$\{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$. Then $r \leq s$.*

Proof 1: Define span$\{\mathbf{y}_1, \cdots, \mathbf{y}_s\} \equiv V$, it follows there exist scalars $c_1, \cdots, c_s$ such that

$$\mathbf{x}_1 = \sum_{i=1}^{s} c_i \mathbf{y}_i. \tag{7.5}$$

Not all of these scalars can equal zero because if this were the case, it would follow that $\mathbf{x}_1 = 0$ and so $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ would not be linearly independent. Indeed, if $\mathbf{x}_1 = 0$, $1\mathbf{x}_1 + \sum_{i=2}^{r} 0\mathbf{x}_i = \mathbf{x}_1 = 0$ and so there would exist a nontrivial linear combination of the vectors $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ which equals zero.

Say $c_k \neq 0$. Then solve 7.5 for $\mathbf{y}_k$ and obtain

$$\mathbf{y}_k \in \text{span}\left(\mathbf{x}_1, \overbrace{\mathbf{y}_1, \cdots, \mathbf{y}_{k-1}, \mathbf{y}_{k+1}, \cdots, \mathbf{y}_s}^{\text{s-1 vectors here}}\right).$$

Define $\{\mathbf{z}_1, \cdots, \mathbf{z}_{s-1}\}$ by

$$\{\mathbf{z}_1, \cdots, \mathbf{z}_{s-1}\} \equiv \{\mathbf{y}_1, \cdots, \mathbf{y}_{k-1}, \mathbf{y}_{k+1}, \cdots, \mathbf{y}_s\}$$

Therefore, $\text{span}\{\mathbf{x}_1, \mathbf{z}_1, \cdots, \mathbf{z}_{s-1}\} = V$ because if $\mathbf{v} \in V$, there exist constants $c_1, \cdots, c_s$ such that

$$\mathbf{v} = \sum_{i=1}^{s-1} c_i \mathbf{z}_i + c_s \mathbf{y}_k.$$

Now replace the $\mathbf{y}_k$ in the above with a linear combination of the vectors, $\{\mathbf{x}_1, \mathbf{z}_1, \cdots, \mathbf{z}_{s-1}\}$ to obtain $\mathbf{v} \in \text{span}\{\mathbf{x}_1, \mathbf{z}_1, \cdots, \mathbf{z}_{s-1}\}$. The vector $\mathbf{y}_k$, in the list $\{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$, has now been replaced with the vector $\mathbf{x}_1$ and the resulting modified list of vectors has the same span as the original list of vectors, $\{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$.

Now suppose that $r > s$ and that $\text{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_l, \mathbf{z}_1, \cdots, \mathbf{z}_p\} = V$ where the vectors, $\mathbf{z}_1, \cdots, \mathbf{z}_p$ are each taken from the set, $\{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$ and $l + p = s$. This has now been done for $l = 1$ above. Then since $r > s$, it follows that $l \leq s < r$ and so $l + 1 \leq r$. Therefore, $\mathbf{x}_{l+1}$ is a vector not in the list, $\{\mathbf{x}_1, \cdots, \mathbf{x}_l\}$ and since $\text{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_l, \mathbf{z}_1, \cdots, \mathbf{z}_p\} = V$ there exist scalars c_i and d_j such that

$$\mathbf{x}_{l+1} = \sum_{i=1}^{l} c_i \mathbf{x}_i + \sum_{j=1}^{p} d_j \mathbf{z}_j. \tag{7.6}$$

Now not all the d_j can equal zero because if this were so, it would follow that $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ would be a linearly dependent set because one of the vectors would equal a linear combination of the others. Therefore, (7.6) can be solved for one of the $\mathbf{z}_i$, say $\mathbf{z}_k$, in terms of $\mathbf{x}_{l+1}$ and the other $\mathbf{z}_i$ and just as in the above argument, replace that $\mathbf{z}_i$ with $\mathbf{x}_{l+1}$ to obtain

$$\text{span}\left(\mathbf{x}_1, \cdots \mathbf{x}_l, \mathbf{x}_{l+1}, \overbrace{\mathbf{z}_1, \cdots \mathbf{z}_{k-1}, \mathbf{z}_{k+1}, \cdots, \mathbf{z}_p}^{\text{p-1 vectors here}}\right) = V.$$

Continue this way, eventually obtaining

$$\text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_s) = V.$$

But then $\mathbf{x}_r \in \text{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_s\}$ contrary to the assumption that $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ is linearly independent. Therefore, $r \leq s$ as claimed.

Proof 2: Let

$$\mathbf{x}_k = \sum_{j=1}^{s} a_{jk}\mathbf{y}_j$$

If $r > s$, then the matrix $A = (a_{jk})$ has more columns than rows. By Corollary 4.3.9 one of these columns is a linear combination of the others. This implies there exist scalars $c_1, \cdots, c_r$, not all zero such that

$$\sum_{k=1}^{r} a_{jk}c_k = 0, \; j = 1, \cdots, r$$

Then

$$\sum_{k=1}^{r} c_k\mathbf{x}_k = \sum_{k=1}^{r} c_k \sum_{j=1}^{s} a_{jk}\mathbf{y}_j = \sum_{j=1}^{s}\left(\sum_{k=1}^{r} c_k a_{jk}\right)\mathbf{y}_j = \mathbf{0}$$

which contradicts the assumption that $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ is linearly independent. Hence $r \leq s$.

Proof 3: Suppose $r > s$. Let $\mathbf{z}_k$ denote a vector of $\{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$. Thus there exists j as small as possible such that

$$\text{span}(\mathbf{y}_1, \cdots, \mathbf{y}_s) = \text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_m, \mathbf{z}_1, \cdots, \mathbf{z}_j)$$

where $m + j = s$. It is given that $m = 0$, corresponding to no vectors of $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ and $j = s$, corresponding to all the $\mathbf{y}_k$ results in the above equation holding. If $j > 0$ then $m < s$ and so

$$\mathbf{x}_{m+1} = \sum_{k=1}^{m} a_k\mathbf{x}_k + \sum_{i=1}^{j} b_i\mathbf{z}_i$$

Not all the b_i can equal 0 and so you can solve for one of them in terms of $\mathbf{x}_{m+1}, \mathbf{x}_m, \cdots, \mathbf{x}_1$, and the other $\mathbf{z}_k$. Therefore, there exists

$$\{\mathbf{z}_1, \cdots, \mathbf{z}_{j-1}\} \subseteq \{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$$

such that

$$\text{span}(\mathbf{y}_1, \cdots, \mathbf{y}_s) = \text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_{m+1}, \mathbf{z}_1, \cdots, \mathbf{z}_{j-1})$$

contradicting the choice of j. Hence $j = 0$ and

$$\text{span}(\mathbf{y}_1, \cdots, \mathbf{y}_s) = \text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_s)$$

It follows that

$$\mathbf{x}_{s+1} \in \text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_s)$$

contrary to the assumption the $\mathbf{x}_k$ are linearly independent. Therefore, $r \leq s$ as claimed. ■

Corollary 7.2.5 *If $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ and $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ are two bases for V, then $m = n$.*

Proof: By Theorem 7.2.4, $m \leq n$ and $n \leq m$. ■

Definition 7.2.6 *A vector space V is of dimension n if it has a basis consisting of n vectors. This is well defined thanks to Corollary 7.2.5. It is always assumed here that $n < \infty$ and in this case, such a vector space is said to be finite dimensional.*

Example 7.2.7 *Consider the polynomials defined on $\mathbb{R}$ of degree no more than 3, denoted here as P_3. Then show that a basis for P_3 is $\{1, x, x^2, x^3\}$. Here x^k symbolizes the function $x \mapsto x^k$.*

It is obvious that the span of the given vectors yields P_3. Why is this set of vectors linearly independent? Suppose

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 = 0$$

where 0 is the zero function which maps everything to 0. Then you could differentiate three times and obtain the following equations

$$\begin{aligned} c_1 + 2c_2 x + 3c_3 x^2 &= 0 \\ 2c_2 + 6c_3 x &= 0 \\ 6c_3 &= 0 \end{aligned}$$

Now this implies $c_3 = 0$. Then from the equations above the bottom one, you find in succession that $c_2 = 0, c_1 = 0, c_0 = 0$.

There is a somewhat interesting theorem about linear independence of smooth functions (those having plenty of derivatives) which I will show now. It is often used in differential equations.

Definition 7.2.8 *Let $f_1, \cdots, f_n$ be smooth functions defined on an interval $[a, b]$. The Wronskian of these functions is defined as follows.*

$$W(f_1, \cdots, f_n)(x) \equiv \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

Note that to get from one row to the next, you just differentiate everything in that row. The notation $f^{(k)}(x)$ denotes the k^{th} derivative.

With this definition, the following is the theorem. The interesting theorem involving the Wronskian has to do with the situation where the functions are solutions of a differential equation. Then much more can be said and it is much more interesting than the following theorem.

Theorem 7.2.9 *Let $\{f_1, \cdots, f_n\}$ be smooth functions defined on $[a, b]$. Then they are linearly independent if there exists some point $t \in [a, b]$ where $W(f_1, \cdots, f_n)(t) \neq 0$.*

Proof: Form the linear combination of these vectors (functions) and suppose it equals 0. Thus

$$a_1 f_1 + a_2 f_2 + \cdots + a_n f_n = 0$$

The question you must answer is whether this requires each a_j to equal zero. If they all must equal 0, then this means these vectors (functions) are independent. This is what it means to be linearly independent.

Differentiate the above equation $n - 1$ times yielding the equations

$$\begin{pmatrix} a_1 f_1 + a_2 f_2 + \cdots + a_n f_n = 0 \\ a_1 f_1' + a_2 f_2' + \cdots + a_n f_n' = 0 \\ \vdots \\ a_1 f_1^{(n-1)} + a_2 f_2^{(n-1)} + \cdots + a_n f_n^{(n-1)} = 0 \end{pmatrix}$$

Now plug in t. Then the above yields

$$\begin{pmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f_1'(t) & f_2'(t) & \cdots & f_n'(t) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Since the determinant of the matrix on the left is assumed to be nonzero, it follows this matrix has an inverse and so the only solution to the above system of equations is to have each $a_k = 0$. ■

Here is a useful lemma.

Lemma 7.2.10 *Suppose* $\mathbf{v} \notin \text{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k)$ *and* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ *is linearly independent. Then* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{v}\}$ *is also linearly independent.*

Proof: Suppose $\sum_{i=1}^{k} c_i \mathbf{u}_i + d\mathbf{v} = 0$. It is required to verify that each $c_i = 0$ and that $d = 0$. But if $d \neq 0$, then you can solve for $\mathbf{v}$ as a linear combination of the vectors, $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$,

$$\mathbf{v} = -\sum_{i=1}^{k} \left(\frac{c_i}{d}\right) \mathbf{u}_i$$

contrary to assumption. Therefore, $d = 0$. But then $\sum_{i=1}^{k} c_i \mathbf{u}_i = 0$ and the linear independence of $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ implies each $c_i = 0$ also. ■

Given a spanning set, you can delete vectors till you end up with a basis. Given a linearly independent set, you can add vectors till you get a basis. This is what the following theorem is about, weeding and planting.

Theorem 7.2.11 *If* $V = \text{span}(\mathbf{u}_1, \cdots, \mathbf{u}_n)$ *then some subset of* $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ *is a basis for* V. *Also, if* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\} \subseteq V$ *is linearly independent and the vector space is finite dimensional, then the set,* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$, *can be enlarged to obtain a basis of* V.

Proof: Let

$$S = \{E \subseteq \{\mathbf{u}_1, \cdots, \mathbf{u}_n\} \text{ such that } \text{span}(E) = V\}.$$

For $E \in S$, let $|E|$ denote the number of elements of E. Let

$$m \equiv \min\{|E| \text{ such that } E \in S\}.$$

Thus there exist vectors

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_m\} \subseteq \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$$

such that

$$\text{span}(\mathbf{v}_1, \cdots, \mathbf{v}_m) = V$$

and m is as small as possible for this to happen. If this set is linearly independent, it follows it is a basis for V and the theorem is proved. On the other hand, if the set is not linearly independent, then there exist scalars

$$c_1, \cdots, c_m$$

such that

$$\mathbf{0} = \sum_{i=1}^{m} c_i \mathbf{v}_i$$

and not all the c_i are equal to zero. Suppose $c_k \neq 0$. Then the vector, $\mathbf{v}_k$ may be solved for in terms of the other vectors. Consequently,

$$V = \text{span}(\mathbf{v}_1, \cdots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \cdots, \mathbf{v}_m)$$

contradicting the definition of m. This proves the first part of the theorem.

To obtain the second part, begin with $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ and suppose a basis for V is

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}.$$

If

$$\operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k) = V,$$

then $k = n$. If not, there exists a vector,

$$\mathbf{u}_{k+1} \notin \operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k).$$

Then by Lemma 7.2.10, $\{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ is also linearly independent. Continue adding vectors in this way until n linearly independent vectors have been obtained. Then

$$\operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_n) = V$$

because if it did not do so, there would exist $\mathbf{u}_{n+1}$ as just described and $\{\mathbf{u}_1, \cdots, \mathbf{u}_{n+1}\}$ would be a linearly independent set of vectors having $n+1$ elements even though $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis. This would contradict Theorem 7.2.4. Therefore, this list is a basis. ■

7.2.3 The Basis Of A Subspace

Every subspace of a finite dimensional vector space is a span of some vectors and in fact it has a basis. This is the content of the next theorem.

Theorem 7.2.12 *Let V be a nonzero subspace of a finite dimensional vector space W of dimension n. Then V has a basis with no more than n vectors.*

Proof: Let $\mathbf{v}_1 \in V$ where $\mathbf{v}_1 \neq 0$. If $\text{span}\{\mathbf{v}_1\} = V$, stop. $\{\mathbf{v}_1\}$ is a basis for V. Otherwise, there exists $\mathbf{v}_2 \in V$ which is not in $\text{span}\{\mathbf{v}_1\}$. By Lemma 7.2.10 $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set of vectors. If $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = V$ stop, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for V. If $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} \neq V$, then there exists $\mathbf{v}_3 \notin \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a larger linearly independent set of vectors. Continuing this way, the process must stop before $n+1$ steps because if not, it would be possible to obtain $n+1$ linearly independent vectors contrary to the exchange theorem, Theorem 7.2.4. ■

7.3 Lots Of Fields

7.3.1 Irreducible Polynomials

I mentioned earlier that most things hold for arbitrary fields. However, I have not bothered to give any examples of other fields. This is the point of this section. It also turns out that showing the algebraic numbers are a field can be understood using vector space concepts and it gives a very convincing application of the abstract theory presented earlier in this chapter.

Here I will give some basic algebra relating to polynomials. This is interesting for its own sake but also provides the basis for constructing many different kinds of fields. The first is the Euclidean algorithm for polynomials.

Definition 7.3.1 *A polynomial is an expression of the form $p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ where as usual λ^0 is defined to equal 1. Two polynomials are said to be equal if their corresponding coefficients are the same. Thus, in particular, $p(\lambda) = 0$ means each of the $a_k = 0$. An element of the field λ is said to be a root of the polynomial if $p(\lambda) = 0$ in the sense that when you plug in λ into the formula and do the indicated operations, you get 0. The degree of a nonzero polynomial is the highest exponent appearing on λ. The degree of the zero polynomial $p(\lambda) = 0$ is not defined. A polynomial of degree n is monic if the coefficient of λ^n is 1. In any case, this coefficient is called the leading coefficient.*

Example 7.3.2 *Consider the polynomial $p(\lambda) = \lambda^2 + \lambda$ where the coefficients are in $\mathbb{Z}_2$. Is this polynomial equal to 0? Not according to the above definition, because its coefficients are not all equal to 0. However, $p(1) = p(0) = 0$ so it sends every element of $\mathbb{Z}_2$ to 0. Note the distinction between saying it sends everything in the field to 0 with having the polynomial be the zero polynomial.*

The fundamental result is the division theorem for polynomials. It is Lemma 1.10.10 on Page 27. We state it here for convenience.

Lemma 7.3.3 *Let $f(\lambda)$ and $g(\lambda) \neq 0$ be polynomials. Then there exists a polynomial, $q(\lambda)$ such that*

$$f(\lambda) = q(\lambda) g(\lambda) + r(\lambda)$$

where the degree of $r(\lambda)$ is less than the degree of $g(\lambda)$ or $r(\lambda) = 0$. These polynomials $q(\lambda)$ and $r(\lambda)$ are unique.

In what follows, the coefficients of polynomials are in $\mathbb{F}$, a field of scalars which is completely arbitrary. Think $\mathbb{R}$ if you need an example.

Definition 7.3.4 *A polynomial f is said to divide a polynomial g if $g(\lambda) = f(\lambda) r(\lambda)$ for some polynomial $r(\lambda)$. Let $\{\phi_i(\lambda)\}$ be a finite set of polynomials. The greatest common divisor will be the* ***monic*** *polynomial $q(\lambda)$ such that $q(\lambda)$ divides each $\phi_i(\lambda)$ and if $p(\lambda)$ divides each $\phi_i(\lambda)$, then $p(\lambda)$ divides $q(\lambda)$. The finite set of polynomials $\{\phi_i\}$ is said to be relatively prime if their greatest common divisor is 1. A polynomial $f(\lambda)$ is irreducible if there is no polynomial with coefficients in $\mathbb{F}$ which divides it except nonzero scalar multiples of $f(\lambda)$ and constants. In other words, it is not possible to write $f(\lambda) = a(\lambda) b(\lambda)$ where each of $a(\lambda), b(\lambda)$ have degree less than the degree of $f(\lambda)$.*

Proposition 7.3.5 *The greatest common divisor is unique.*

Proof: Suppose both $q(\lambda)$ and $q'(\lambda)$ work. Then $q(\lambda)$ divides $q'(\lambda)$ and the other way around and so

$$q'(\lambda) = q(\lambda) l(\lambda),\ q(\lambda) = l'(\lambda) q'(\lambda)$$

Therefore, the two must have the same degree. Hence $l'(\lambda), l(\lambda)$ are both constants. However, this constant must be 1 because both $q(\lambda)$ and $q'(\lambda)$ are monic. ■

Theorem 7.3.6 *Let $\psi(\lambda)$ be the greatest common divisor of $\{\phi_i(\lambda)\}$, not all of which are zero polynomials. Then there exist polynomials $r_i(\lambda)$ such that*

$$\psi(\lambda) = \sum_{i=1}^{p} r_i(\lambda) \phi_i(\lambda).$$

Furthermore, $\psi(\lambda)$ is the monic polynomial of smallest degree which can be written in the above form.

Proof: Let S denote the set of monic polynomials which are of the form

$$\sum_{i=1}^{p} r_i(\lambda) \phi_i(\lambda)$$

where $r_i(\lambda)$ is a polynomial. Then $S \neq \emptyset$ because some $\phi_i(\lambda) \neq 0$. Then let the r_i be chosen such that the degree of the expression $\sum_{i=1}^{p} r_i(\lambda) \phi_i(\lambda)$ is as small as possible. Letting $\psi(\lambda)$ equal this sum, it remains to verify it is the greatest common divisor. First, does it divide each $\phi_i(\lambda)$? Suppose it fails to divide $\phi_1(\lambda)$. Then by Lemma 7.3.3

$$\phi_1(\lambda) = \psi(\lambda) l(\lambda) + r(\lambda)$$

where degree of $r(\lambda)$ is less than that of $\psi(\lambda)$. Then dividing $r(\lambda)$ by the leading coefficient if necessary and denoting the result by $\psi_1(\lambda)$, it follows the degree of $\psi_1(\lambda)$ is less than the degree of $\psi(\lambda)$ and $\psi_1(\lambda)$ equals

$$\psi_1(\lambda) = (\phi_1(\lambda) - \psi(\lambda) l(\lambda)) a$$

$$\begin{aligned} &= \left(\phi_1(\lambda) - \sum_{i=1}^{p} r_i(\lambda) \phi_i(\lambda) l(\lambda)\right) a \\ &= \left((1 - r_1(\lambda)) \phi_1(\lambda) + \sum_{i=2}^{p} (-r_i(\lambda) l(\lambda)) \phi_i(\lambda)\right) a \end{aligned}$$

for a suitable $a \in \mathbb{F}$. This is one of the polynomials in S. Therefore, $\psi(\lambda)$ does not have the smallest degree after all because the degree of $\psi_1(\lambda)$ is smaller. This is a contradiction. Therefore, $\psi(\lambda)$ divides $\phi_1(\lambda)$. Similarly it divides all the other $\phi_i(\lambda)$.

If $p(\lambda)$ divides all the $\phi_i(\lambda)$, then it divides $\psi(\lambda)$ because of the formula for $\psi(\lambda)$ which equals $\sum_{i=1}^{p} r_i(\lambda) \phi_i(\lambda)$. ■

Lemma 7.3.7 *Suppose $\phi(\lambda)$ and $\psi(\lambda)$ are monic polynomials which are irreducible and not equal. Then they are relatively prime.*

Proof: Suppose $\eta(\lambda)$ is a nonconstant polynomial. If $\eta(\lambda)$ divides $\phi(\lambda)$, then since $\phi(\lambda)$ is irreducible, $\eta(\lambda)$ equals $a\phi(\lambda)$ for some $a \in \mathbb{F}$. If $\eta(\lambda)$ divides $\psi(\lambda)$ then it must be of the form $b\psi(\lambda)$ for some $b \in \mathbb{F}$ and so it follows

$$\psi(\lambda) = \frac{a}{b} \phi(\lambda)$$

but both $\psi(\lambda)$ and $\phi(\lambda)$ are monic polynomials which implies $a = b$ and so $\psi(\lambda) = \phi(\lambda)$. This is assumed not to happen. It follows the only polynomials which divide both $\psi(\lambda)$ and $\phi(\lambda)$ are constants and so the two polynomials are relatively prime. Thus a polynomial which divides them both must be a constant, and if it is monic, then it must be 1. Thus 1 is the greatest common divisor. ■

Lemma 7.3.8 *Let $\psi(\lambda)$ be an irreducible monic polynomial not equal to 1 which divides*

$$\prod_{i=1}^{p} \phi_i(\lambda)^{k_i}, \ k_i \text{ a positive integer,}$$

where each $\phi_i(\lambda)$ is an irreducible monic polynomial not equal to 1. Then $\psi(\lambda)$ equals some $\phi_i(\lambda)$.

Proof : Say $\psi(\lambda)\, l(\lambda) = \prod_{i=1}^{p} \phi_i(\lambda)^{k_i}$. Suppose $\psi(\lambda) \neq \phi_i(\lambda)$ for all i. Then by Lemma 7.3.7, there exist polynomials $m_i(\lambda), n_i(\lambda)$ such that

$$\begin{aligned} 1 &= \psi(\lambda) m_i(\lambda) + \phi_i(\lambda) n_i(\lambda) \\ \phi_i(\lambda) n_i(\lambda) &= 1 - \psi(\lambda) m_i(\lambda) \end{aligned}$$

Hence,

$$\begin{aligned} \psi(\lambda) n(\lambda) &\equiv \psi(\lambda) l(\lambda) \prod_{i=1}^{p} n_i(\lambda)^{k_i} = \prod_{i=1}^{p} (n_i(\lambda) \phi_i(\lambda))^{k_i} \\ &= \prod_{i=1}^{p} (1 - \psi(\lambda) m_i(\lambda))^{k_i} = 1 + g(\lambda) \psi(\lambda) \end{aligned}$$

for a polynomial $g(\lambda)$. Thus

$$1 = \psi(\lambda)(n(\lambda) - g(\lambda))$$

which is impossible because $\psi(\lambda)$ is not equal to 1. ■

Now here is a simple lemma about canceling monic polynomials.

Lemma 7.3.9 *Suppose $p(\lambda)$ is a monic polynomial and $q(\lambda)$ is a polynomial such that*

$$p(\lambda) q(\lambda) = 0.$$

Then $q(\lambda) = 0$. Also if

$$p(\lambda) q_1(\lambda) = p(\lambda) q_2(\lambda)$$

then $q_1(\lambda) = q_2(\lambda)$.

Proof: Let

$$p(\lambda) = \sum_{j=1}^{k} p_j \lambda^j,\ q(\lambda) = \sum_{i=1}^{n} q_i \lambda^i,\ p_k = 1.$$

Then the product equals

$$\sum_{j=1}^{k} \sum_{i=1}^{n} p_j q_i \lambda^{i+j}.$$

Then look at those terms involving λ^{k+n}. This is $p_k q_n \lambda^{k+n}$ and is given to be 0. Since $p_k = 1$, it follows $q_n = 0$. Thus

$$\sum_{j=1}^{k} \sum_{i=1}^{n-1} p_j q_i \lambda^{i+j} = 0.$$

Then consider the term involving λ^{n-1+k} and conclude that since $p_k = 1$, it follows $q_{n-1} = 0$. Continuing this way, each $q_i = 0$. This proves the first part. The second follows from

$$p(\lambda)(q_1(\lambda) - q_2(\lambda)) = 0. \ ■$$

The following is the analog of the fundamental theorem of arithmetic for polynomials.

Theorem 7.3.10 *Let $f(\lambda)$ be a nonconstant polynomial with coefficients in $\mathbb{F}$. Then there is some $a \in \mathbb{F}$ such that $f(\lambda) = a \prod_{i=1}^{n} \phi_i(\lambda)$ where $\phi_i(\lambda)$ is an irreducible nonconstant monic polynomial and repeats are allowed. Furthermore, this factorization is unique in the sense that any two of these factorizations have the same nonconstant factors in the product, possibly in different order and the same constant a.*

Proof: That such a factorization exists is obvious. If $f(\lambda)$ is irreducible, you are done. Factor out the leading coefficient. If not, then $f(\lambda) = a\phi_1(\lambda)\phi_2(\lambda)$ where these are monic polynomials. Continue doing this with the ϕ_i and eventually arrive at a factorization of the desired form.

It remains to argue the factorization is unique except for order of the factors. Suppose

$$a \prod_{i=1}^{n} \phi_i(\lambda) = b \prod_{i=1}^{m} \psi_i(\lambda)$$

where the $\phi_i(\lambda)$ and the $\psi_i(\lambda)$ are all irreducible monic nonconstant polynomials and $a, b \in \mathbb{F}$. If $n > m$, then by Lemma 7.3.8, each $\psi_i(\lambda)$ equals one of the $\phi_j(\lambda)$. By the above cancellation lemma, Lemma 7.3.9, you can cancel all these $\psi_i(\lambda)$ with appropriate $\phi_j(\lambda)$ and obtain a contradiction because the resulting polynomials on either side would have different degrees. Similarly, it cannot happen that $n < m$. It follows $n = m$ and the two products consist of the same polynomials. Then it follows $a = b$. ■

The following corollary will be well used. This corollary seems rather believable but does require a proof.

Corollary 7.3.11 *Let $q(\lambda) = \prod_{i=1}^{p} \phi_i(\lambda)^{k_i}$ where the k_i are positive integers and the $\phi_i(\lambda)$ are irreducible monic polynomials. Suppose also that $p(\lambda)$ is a monic polynomial which divides $q(\lambda)$. Then*

$$p(\lambda) = \prod_{i=1}^{p} \phi_i(\lambda)^{r_i}$$

where r_i is a nonnegative integer no larger than k_i.

Proof: Using Theorem 7.3.10, let $p(\lambda) = b\prod_{i=1}^{s} \psi_i(\lambda)^{r_i}$ where the $\psi_i(\lambda)$ are each irreducible and monic and $b \in \mathbb{F}$. Since $p(\lambda)$ is monic, $b = 1$. Then there exists a polynomial $g(\lambda)$ such that

$$p(\lambda) g(\lambda) = g(\lambda) \prod_{i=1}^{s} \psi_i(\lambda)^{r_i} = \prod_{i=1}^{p} \phi_i(\lambda)^{k_i}$$

Hence $g(\lambda)$ must be monic. Therefore,

$$p(\lambda) g(\lambda) = \overbrace{\prod_{i=1}^{s} \psi_i(\lambda)^{r_i}}^{p(\lambda)} \prod_{j=1}^{l} \eta_j(\lambda) = \prod_{i=1}^{p} \phi_i(\lambda)^{k_i}$$

for η_j monic and irreducible. By uniqueness, each ψ_i equals one of the $\phi_j(\lambda)$ and the same holding true of the $\eta_i(\lambda)$. Therefore, $p(\lambda)$ is of the desired form. ■

7.3.2 Polynomials And Fields

When you have a polynomial like $x^2 - 3$ which has no rational roots, it turns out you can enlarge the field of rational numbers to obtain a larger field such that this polynomial does have roots in this larger field. I am going to discuss a systematic way to do this. It will turn out that for any polynomial with coefficients in any field, there always exists a possibly larger field such that the polynomial has roots in this larger field. This book has mainly featured the field of real or complex numbers but this procedure will show how to obtain many other fields which could be used in most of what was presented earlier in the book. Here is an important idea concerning equivalence relations which I hope is familiar.

Definition 7.3.12 *Let S be a set. The symbol, $\sim$ is called an equivalence relation on S if it satisfies the following axioms.*

1. *$x \sim x$ for all $x \in S$. (Reflexive)*
2. *If $x \sim y$ then $y \sim x$. (Symmetric)*
3. *If $x \sim y$ and $y \sim z$, then $x \sim z$. (Transitive)*

Definition 7.3.13 *$[x]$ denotes the set of all elements of S which are equivalent to x and $[x]$ is called the equivalence class determined by x or just the equivalence class of x.*

Also recall the notion of equivalence classes.

Theorem 7.3.14 *Let $\sim$ be an equivalence class defined on a set, S and let $\mathcal{H}$ denote the set of equivalence classes. Then if $[x]$ and $[y]$ are two of these equivalence classes, either $x \sim y$ and $[x] = [y]$ or it is not true that $x \sim y$ and $[x] \cap [y] = \emptyset$.*

Definition 7.3.15 *Let $\mathbb{F}$ be a field, for example the rational numbers, and denote by $\mathbb{F}[x]$ the polynomials having coefficients in $\mathbb{F}$. Suppose $p(x)$ is a polynomial. Let $a(x) \sim b(x)$ ($a(x)$ is similar to $b(x)$) when*

$$a(x) - b(x) = k(x) p(x)$$

for some polynomial $k(x)$.

Proposition 7.3.16 *In the above definition, $\sim$ is an equivalence relation.*

Proof: First of all, note that $a(x) \sim a(x)$ because their difference equals $0p(x)$. If $a(x) \sim b(x)$, then $a(x) - b(x) = k(x)p(x)$ for some $k(x)$. But then $b(x) - a(x) = -k(x)p(x)$ and so $b(x) \sim a(x)$. Next suppose $a(x) \sim b(x)$ and $b(x) \sim c(x)$. Then $a(x) - b(x) = k(x)p(x)$ for some polynomial $k(x)$ and also $b(x) - c(x) = l(x)p(x)$ for some polynomial $l(x)$. Then

$$a(x) - c(x) = a(x) - b(x) + b(x) - c(x)$$

$$= k(x)p(x) + l(x)p(x) = (l(x) + k(x))p(x)$$

and so $a(x) \sim c(x)$ and this shows the transitive law. ■

With this proposition, here is another definition which essentially describes the elements of the new field. It will eventually be necessary to assume the polynomial $p(x)$ in the above definition is irreducible so I will begin assuming this.

Definition 7.3.17 *Let $\mathbb{F}$ be a field and let $p(x) \in \mathbb{F}[x]$ be a monic irreducible polynomial of degree greater than 0. Thus there is no polynomial having coefficients in $\mathbb{F}$ which divides $p(x)$ except for itself and constants, and its leading coefficient is 1. For the similarity relation defined in Definition 7.3.15, define the following operations on the equivalence classes. $[a(x)]$ is an equivalence class means that it is the set of all polynomials which are similar to $a(x)$.*

$$\begin{aligned} [a(x)] + [b(x)] &\equiv [a(x) + b(x)] \\ [a(x)][b(x)] &\equiv [a(x)b(x)] \end{aligned}$$

This collection of equivalence classes is sometimes denoted by $\mathbb{F}[x]/(p(x))$.

Proposition 7.3.18 *In the situation of Definition 7.3.17 where $p(x)$ is a monic irreducible polynomial, the following are valid.*

1. *$p(x)$ and $q(x)$ are relatively prime for any $q(x) \in \mathbb{F}[x]$ which is not a multiple of $p(x)$.*

2. *The definitions of addition and multiplication are well defined.*

3. *If $a, b \in \mathbb{F}$ and $[a] = [b]$, then $a = b$. Thus $\mathbb{F}$ can be considered a subset of $\mathbb{F}[x] / (p(x))$.*

4. *$\mathbb{F}[x] / (p(x))$ is a field in which the polynomial $p(x)$ has a root.*

5. *$\mathbb{F}[x] / (p(x))$ is a vector space with field of scalars $\mathbb{F}$ and its dimension is m where m is the degree of the irreducible polynomial $p(x)$.*

Proof: First consider the claim about $p(x), q(x)$ being relatively prime. If $\psi(x)$ is the greatest common divisor, it follows $\psi(x)$ is either equal to $p(x)$ or 1. If it is $p(x)$, then $q(x)$ is a multiple of $p(x)$ which does not happen. If it is 1, then by definition, the two polynomials are relatively prime.

To show the operations are well defined, suppose

$$[a(x)] = [a'(x)], [b(x)] = [b'(x)]$$

It is necessary to show

$$[a(x) + b(x)] = [a'(x) + b'(x)]$$
$$[a(x) b(x)] = [a'(x) b'(x)]$$

Consider the second of the two.

$$\begin{aligned} & a'(x) b'(x) - a(x) b(x) \\ = \; & a'(x) b'(x) - a(x) b'(x) + a(x) b'(x) - a(x) b(x) \\ = \; & b'(x) (a'(x) - a(x)) + a(x) (b'(x) - b(x)) \end{aligned}$$

Now by assumption $(a'(x) - a(x))$ is a multiple of $p(x)$ as is $(b'(x) - b(x))$, so the above is a multiple of $p(x)$ and by definition this shows $[a(x) b(x)] = [a'(x) b'(x)]$. The case for addition is similar.

Now suppose $[a] = [b]$. This means $a - b = k(x) p(x)$ for some polynomial $k(x)$. Then $k(x)$ must equal 0 since otherwise the two polynomials $a - b$ and $k(x) p(x)$ could not be equal because they would have different degree.

It is clear that the axioms of a field are satisfied except for the one which says that non zero elements of the field have a multiplicative inverse. Let $[q(x)] \in \mathbb{F}[x] / (p(x))$ where $[q(x)] \neq [0]$. Then $q(x)$ is not a multiple of $p(x)$ and so by the first part, $q(x), p(x)$ are relatively prime. Thus there exist $n(x), m(x)$ such that

$$1 = n(x) q(x) + m(x) p(x)$$

Hence

$$[1] = [1 - n(x) p(x)] = [n(x) q(x)] = [n(x)] [q(x)]$$

which shows that $[q(x)]^{-1} = [n(x)]$. Thus this is a field. The polynomial has a root in this field because if

$$p(x) = x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0,$$
$$[0] = [p(x)] = [x]^m + [a_{m-1}] [x]^{m-1} + \cdots + [a_1] [x] + [a_0]$$

Thus $[x]$ is a root of this polynomial in the field $\mathbb{F}[x] / (p(x))$.

Consider the last claim. Let $f(x) \in \mathbb{F}[x] / (p(x))$. Thus $[f(x)]$ is a typical thing in $\mathbb{F}[x] / (p(x))$. Then from the division algorithm,

$$f(x) = p(x) q(x) + r(x)$$

where $r(x)$ is either 0 or has degree less than the degree of $p(x)$. Thus

$$[r(x)] = [f(x) - p(x) q(x)] = [f(x)]$$

but clearly $[r(x)] \in \operatorname{span}\left([1], \cdots, [x]^{m-1}\right)$. Thus $\operatorname{span}\left([1], \cdots, [x]^{m-1}\right) = \mathbb{F}[x] / (p(x))$. Then $\left\{[1], \cdots, [x]^{m-1}\right\}$ is a basis if these vectors are linearly independent. Suppose then that

$$\sum_{i=0}^{m-1} c_i [x]^i = \left[\sum_{i=0}^{m-1} c_i x^i\right] = 0$$

Then you would need to have $p(x) / \sum_{i=0}^{m-1} c_i x^i$ which is impossible unless each $c_i = 0$ because $p(x)$ has degree m. ■

From the above theorem, it makes perfect sense to write b rather than $[b]$ if $b \in \mathbb{F}$. Then with this convention,

$$[b\phi(x)] = [b][\phi(x)] = b[\phi(x)].$$

This shows how to enlarge a field to get a new one in which the polynomial has a root. By using a succession of such enlargements, called field extensions, there will exist a field in which the given polynomial can be factored into a product of polynomials having degree one. The field you obtain in this process of enlarging in which the given polynomial factors in terms of linear factors is called a splitting field.

Remark 7.3.19 *The polynomials consisting of all polynomial multiples of $p(x)$, denoted by $(p(x))$ is called an ideal. An ideal I is a subset of the commutative ring (Here the ring is $\mathbb{F}[x]$.) with unity consisting of all polynomials which is itself a ring and which has the property that whenever $f(x) \in \mathbb{F}[x]$, and $g(x) \in I$, $f(x)g(x) \in I$. In this case, you could argue that $(p(x))$ is an ideal and that the only ideal containing it is itself or the entire ring $\mathbb{F}[x]$. This is called a maximal ideal.*

Example 7.3.20 *The polynomial $x^2 - 2$ is irreducible in $\mathbb{Q}[x]$. This is because if $x^2 - 2 = p(x)q(x)$ where $p(x), q(x)$ both have degree less than 2, then they both have degree 1. Hence you would have $x^2 - 2 = (x + a)(x + b)$ which requires that $a + b = 0$ so this factorization is of the form $(x - a)(x + a)$ and now you need to have $a = \sqrt{2} \notin \mathbb{Q}$. Now $\mathbb{Q}[x] / \left(x^2 - 2\right)$ is of the form $a + b[x]$ where $a, b \in \mathbb{Q}$ and $[x]^2 - 2 = 0$. Thus one can regard $[x]$ as $\sqrt{2}$. $\mathbb{Q}[x] / \left(x^2 - 2\right)$ is of the form $a + b\sqrt{2}$.*

In the above example, $\left[x^2 + x\right]$ is not zero because it is not a multiple of $x^2 - 2$. What is $\left[x^2 + x\right]^{-1}$? You know that the two polynomials are relatively prime and so there exists $n(x), m(x)$ such that

$$1 = n(x)\left(x^2 - 2\right) + m(x)\left(x^2 + x\right)$$

Thus $[m(x)] = \left[x^2 + x\right]^{-1}$. How could you find these polynomials? First of all, it suffices to consider only $n(x)$ and $m(x)$ having degree less than 2.

$$1 = (ax + b)\left(x^2 - 2\right) + (cx + d)\left(x^2 + x\right)$$

$$1 = ax^3 - 2b + bx^2 + cx^2 + cx^3 + dx^2 - 2ax + dx$$

Now you solve the resulting system of equations.

$$a = \frac{1}{2}, b = -\frac{1}{2}, c = -\frac{1}{2}, d = 1$$

Then the desired inverse is $\left[-\frac{1}{2}x + 1\right]$. To check,

$$\left(-\frac{1}{2}x + 1\right)\left(x^2 + x\right) - 1 = -\frac{1}{2}(x - 1)\left(x^2 - 2\right)$$

Thus $\left[-\frac{1}{2}x + 1\right]\left[x^2 + x\right] - [1] = [0]$.

The above is an example of something general described in the following definition.

Definition 7.3.21 *Let $F \subseteq K$ be two fields. Then clearly K is also a vector space over F. Then also, K is called a finite field extension of F if the dimension of this vector space, denoted by $[K : F]$ is finite.*

There are some easy things to observe about this.

Proposition 7.3.22 *Let $F \subseteq K \subseteq L$ be fields. Then $[L : F] = [L : K] [K : F]$.*

Proof: Let $\{l_i\}_{i=1}^n$ be a basis for L over K and let $\{k_j\}_{j=1}^m$ be a basis of K over F. Then if $l \in L$, there exist unique scalars x_i in K such that

$$l = \sum_{i=1}^{n} x_i l_i$$

Now $x_i \in K$ so there exist f_{ji} such that

$$x_i = \sum_{j=1}^{m} f_{ji} k_j$$

Then it follows that

$$l = \sum_{i=1}^{n} \sum_{j=1}^{m} f_{ji} k_j l_i$$

It follows that $\{k_j l_i\}$ is a spanning set. If

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ji} k_j l_i = 0$$

Then, since the l_i are independent, it follows that

$$\sum_{j=1}^{m} f_{ji} k_j = 0$$

and since $\{k_j\}$ is independent, each $f_{ji} = 0$ for each j for a given arbitrary i. Therefore, $\{k_j l_i\}$ is a basis. ■

Note that if $p(x)$ were not irreducible, then you could find a field extension $\mathbb{G}$ containing a root of $p(x)$ such that $[\mathbb{G} : \mathbb{F}] \leq n$. You could do this by working with an irreducible factor of $p(x)$.

Theorem 7.3.23 *Let $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial with coefficients in a field of scalars $\mathbb{F}$. There exists a larger field $\mathbb{G}$ and $\{z_1, \cdots, z_n\}$ contained in $\mathbb{G}$, listed according to multiplicity, such that*

$$p(x) = \prod_{i=1}^{n} (x - z_i)$$

This larger field is called a splitting field. Furthermore,

$$[\mathbb{G} : \mathbb{F}] \leq n!$$

Proof: From Proposition 7.3.18, there exists a field $\mathbb{F}_1$ such that $p(x)$ has a root, z_1 $(= [x]$ if p is irreducible.) Then by the Euclidean algorithm

$$p(x) = (x - z_1) q_1(x) + r$$

where $r \in \mathbb{F}_1$. Since $p(z_1) = 0$, this requires $r = 0$. Now do the same for $q_1(x)$ that was done for $p(x)$, enlarging the field to $\mathbb{F}_2$ if necessary, such that in this new field

$$q_1(x) = (x - z_2) q_2(x).$$

and so

$$p(x) = (x - z_1)(x - z_2) q_2(x)$$

After n such extensions, you will have obtained the necessary field $\mathbb{G}$.

Finally consider the claim about dimension. By Proposition 7.3.18, there is a larger field $\mathbb{G}_1$ such that $p(x)$ has a root a_1 in $\mathbb{G}_1$ and $[\mathbb{G}_1 : \mathbb{F}] \leq n$. Then

$$p(x) = (x - a_1) q(x)$$

Continue this way until the polynomial equals the product of linear factors. Then by Proposition 7.3.22 applied multiple times, $[\mathbb{G} : \mathbb{F}] \leq n!$. ■

Example 7.3.24 *The polynomial $x^2 + 1$ is irreducible in $\mathbb{R}[x]$, polynomials having real coefficients. To see this is the case, suppose $\psi(x)$ divides $x^2 + 1$. Then*

$$x^2 + 1 = \psi(x) q(x)$$

If the degree of $\psi(x)$ is less than 2, then it must be either a constant or of the form $ax + b$. In the latter case, $-b/a$ must be a zero of the right side, hence of the left but $x^2 + 1$ has no real zeros. Therefore, the degree of $\psi(x)$ must be two and $q(x)$ must be a constant. Thus the only polynomial which divides $x^2 + 1$ are constants and multiples of $x^2 + 1$. Therefore, this shows $x^2 + 1$ is irreducible. Find the inverse of $\left[x^2 + x + 1\right]$ in the space of equivalence classes, $\mathbb{R}/\left(x^2 + 1\right)$.

You can solve this with partial fractions.

$$\frac{1}{(x^2 + 1)(x^2 + x + 1)} = -\frac{x}{x^2 + 1} + \frac{x + 1}{x^2 + x + 1}$$

and so

$$1 = (-x)\left(x^2 + x + 1\right) + (x + 1)\left(x^2 + 1\right)$$

which implies

$$1 \sim (-x)\left(x^2 + x + 1\right)$$

and so the inverse is $[-x]$.

The following proposition is interesting. It was essentially proved above but to emphasize it, here it is again.

Proposition 7.3.25 *Suppose $p(x) \in \mathbb{F}[x]$ is irreducible and has degree n. Then every element of $\mathbb{G} = \mathbb{F}[x]/(p(x))$ is of the form $[0]$ or $[r(x)]$ where the degree of $r(x)$ is less than n.*

Proof: This follows right away from the Euclidean algorithm for polynomials. If $k(x)$ has degree larger than $n-1$, then

$$k(x) = q(x)\, p(x) + r(x)$$

where $r(x)$ is either equal to 0 or has degree less than n. Hence

$$[k(x)] = [r(x)]. \blacksquare$$

Example 7.3.26 *In the situation of the above example, find $[ax+b]^{-1}$ assuming $a^2 + b^2 \neq 0$. Note this includes all cases of interest thanks to the above proposition.*

You can do it with partial fractions as above.

$$\frac{1}{(x^2+1)(ax+b)} = \frac{b - ax}{(a^2+b^2)(x^2+1)} + \frac{a^2}{(a^2+b^2)(ax+b)}$$

and so

$$1 = \frac{1}{a^2+b^2}(b-ax)(ax+b) + \frac{a^2}{(a^2+b^2)}\left(x^2+1\right)$$

Thus

$$\frac{1}{a^2+b^2}(b-ax)(ax+b) \sim 1$$

and so

$$[ax+b]^{-1} = \frac{[(b-ax)]}{a^2+b^2} = \frac{b - a[x]}{a^2+b^2}$$

You might find it interesting to recall that $(ai+b)^{-1} = \frac{b-ai}{a^2+b^2}$.

7.3.3 The Algebraic Numbers

Each polynomial having coefficients in a field $\mathbb{F}$ has a splitting field. Consider the case of all polynomials $p(x)$ having coefficients in a field $\mathbb{F} \subseteq \mathbb{G}$ and consider all roots which are also in $\mathbb{G}$. The theory of vector spaces is very useful in the study of these algebraic numbers. Here is a definition.

Definition 7.3.27 *The algebraic numbers $\mathbb{A}$ are those numbers which are in $\mathbb{G}$ and also roots of some polynomial $p(x)$ having coefficients in $\mathbb{F}$. The minimal polynomial of $a \in \mathbb{A}$ is defined to be the monic polynomial $p(x)$ having smallest degree such that $p(a) = 0$.*

The next theorem is on the uniqueness of the minimal polynomial.

Theorem 7.3.28 *Let $a \in \mathbb{A}$. Then there exists a unique monic irreducible polynomial $p(x)$ having coefficients in $\mathbb{F}$ such that $p(a) = 0$. This polynomial is the minimal polynomial.*

Proof: Let $p(x)$ be a monic polynomial having smallest degree such that $p(a) = 0$. Then $p(x)$ is irreducible because if not, there would exist a polynomial having smaller degree which has a as a root. Now suppose $q(x)$ is monic with smallest degree such that $q(a) = 0$. Then

$$q(x) = p(x)\, l(x) + r(x)$$

where if $r(x) \neq 0$, then it has smaller degree than $p(x)$. But in this case, the equation implies $r(a) = 0$ which contradicts the choice of $p(x)$. Hence $r(x) = 0$ and so, since $q(x)$ has smallest degree, $l(x) = 1$ showing that $p(x) = q(x)$. $\blacksquare$

Definition 7.3.29 *For a an algebraic number, let* $\deg(a)$ *denote the degree of the minimal polynomial of* a.

Also, here is another definition.

Definition 7.3.30 *Let* $a_1, \cdots, a_m$ *be in* $\mathbb{A}$. *A polynomial in* $\{a_1, \cdots, a_m\}$ *will be an expression of the form*

$$\sum_{k_1 \cdots k_n} a_{k_1 \cdots k_n} a_1^{k_1} \cdots a_n^{k_n}$$

where the $a_{k_1 \cdots k_n}$ *are in* $\mathbb{F}$, *each* k_j *is a nonnegative integer, and all but finitely many of the* $a_{k_1 \cdots k_n}$ *equal zero. The collection of such polynomials will be denoted by*

$$\mathbb{F}[a_1, \cdots, a_m].$$

Now notice that for a an algebraic number, $\mathbb{F}[a]$ is a vector space with field of scalars $\mathbb{F}$. Similarly, for $\{a_1, \cdots, a_m\}$ algebraic numbers, $\mathbb{F}[a_1, \cdots, a_m]$ is a vector space with field of scalars $\mathbb{F}$. The following fundamental proposition is important.

Proposition 7.3.31 *Let* $\{a_1, \cdots, a_m\}$ *be algebraic numbers. Then*

$$\dim \mathbb{F}[a_1, \cdots, a_m] \leq \prod_{j=1}^{m} \deg(a_j)$$

and for an algebraic number a,

$$\dim \mathbb{F}[a] = \deg(a)$$

Every element of $\mathbb{F}[a_1, \cdots, a_m]$ *is in* $\mathbb{A}$ *and* $\mathbb{F}[a_1, \cdots, a_m]$ *is a field.*

Proof: Let the minimal polynomial of a be

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

If $q(a) \in \mathbb{F}[a]$, then

$$q(x) = p(x) l(x) + r(x)$$

where $r(x)$ has degree less than the degree of $p(x)$ if it is not zero. Hence $q(a) = r(a)$. Thus $\mathbb{F}[a]$ is spanned by

$$\{1, a, a^2, \cdots, a^{n-1}\}$$

Since $p(x)$ has smallest degree of all polynomials which have a as a root, the above set is also linearly independent. This proves the second claim.

Now consider the first claim. By definition, $\mathbb{F}[a_1, \cdots, a_m]$ is obtained from all linear combinations of products of $\left\{a_1^{k_1}, a_2^{k_2}, \cdots, a_n^{k_n}\right\}$ where the k_i are nonnegative integers. From the first part, it suffices to consider only $k_j \leq \deg(a_j)$. Therefore, there exists a spanning set for $\mathbb{F}[a_1, \cdots, a_m]$ which has

$$\prod_{i=1}^{m} \deg(a_i)$$

entries. By Theorem 7.2.4 this proves the first claim.

Finally consider the last claim. Let $g(a_1, \cdots, a_m)$ be a polynomial in $\{a_1, \cdots, a_m\}$ in $\mathbb{F}[a_1, \cdots, a_m]$. Since

$$\dim \mathbb{F}[a_1, \cdots, a_m] \equiv p \leq \prod_{j=1}^{m} \deg(a_j) < \infty,$$

it follows

$$1, g(a_1, \cdots, a_m), g(a_1, \cdots, a_m)^2, \cdots, g(a_1, \cdots, a_m)^p$$

are dependent. It follows $g(a_1, \cdots, a_m)$ is the root of some polynomial having coefficients in $\mathbb{F}$. Thus everything in $\mathbb{F}[a_1, \cdots, a_m]$ is algebraic. Why is $\mathbb{F}[a_1, \cdots, a_m]$ a field? Let $g(a_1, \cdots, a_m)$ be as just mentioned. Then it has a minimal polynomial,

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

where the $a_i \in \mathbb{F}$. Then $a_0 \neq 0$ or else the polynomial would not be minimal. Therefore,

$$g(a_1, \cdots, a_m)\left(g(a_1, \cdots, a_m)^{q-1} + a_{q-1}g(a_1, \cdots, a_m)^{q-2} + \cdots + a_1\right) = -a_0$$

and so the multiplicative inverse for $g(a_1, \cdots, a_m)$ is

$$\frac{g(a_1, \cdots, a_m)^{q-1} + a_{q-1}g(a_1, \cdots, a_m)^{q-2} + \cdots + a_1}{-a_0} \in \mathbb{F}[a_1, \cdots, a_m].$$

The other axioms of a field are obvious. ■

Now from this proposition, it is easy to obtain the following interesting result about the algebraic numbers.

Theorem 7.3.32 *The algebraic numbers $\mathbb{A}$, those roots of polynomials in $\mathbb{F}[x]$ which are in $\mathbb{G}$, are a field.*

Proof: By definition, each $a \in \mathbb{A}$ has a minimal polynomial. Let $a \neq 0$ be an algebraic number and let $p(x)$ be its minimal polynomial. Then $p(x)$ is of the form

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

where $a_0 \neq 0$. Otherwise $p(x)$ would not have minimal degree. Then plugging in a yields

$$a\frac{\left(a^{n-1} + a_{n-1}a^{n-2} + \cdots + a_1\right)(-1)}{a_0} = 1.$$

and so $a^{-1} = \frac{\left(a^{n-1}+a_{n-1}a^{n-2}+\cdots+a_1\right)(-1)}{a_0} \in \mathbb{F}[a]$. By the proposition, every element of $\mathbb{F}[a]$ is in $\mathbb{A}$ and this shows that for every nonzero element of $\mathbb{A}$, its inverse is also in $\mathbb{A}$. What about products and sums of things in $\mathbb{A}$? Are they still in $\mathbb{A}$? Yes. If $a, b \in \mathbb{A}$, then both $a+b$ and $ab \in \mathbb{F}[a,b]$ and from the proposition, each element of $\mathbb{F}[a,b]$ is in $\mathbb{A}$. ■

A typical example of what is of interest here is when the field $\mathbb{F}$ of scalars is $\mathbb{Q}$, the rational numbers and the field $\mathbb{G}$ is $\mathbb{R}$. However, you can certainly conceive of many other examples by considering the integers mod a prime, for example (See Problem 34 on Page 214 for example.) or any of the fields which occur as field extensions in the above.

There is a very interesting thing about $\mathbb{F}[a_1 \cdots a_n]$ in the case where $\mathbb{F}$ is infinite which says that there exists a single algebraic γ such that $\mathbb{F}[a_1 \cdots a_n] = \mathbb{F}[\gamma]$. In other words, every field extension of this sort is a simple field extension. I found this fact in an early version of [5].

Proposition 7.3.33 *There exists γ such that $\mathbb{F}[a_1 \cdots a_n] = \mathbb{F}[\gamma]$.*

Proof: To begin with, consider $\mathbb{F}[\alpha, \beta]$. Let $\gamma = \alpha + \lambda\beta$. Then by Proposition 7.3.31 γ is an algebraic number and it is also clear

$$\mathbb{F}[\gamma] \subseteq \mathbb{F}[\alpha, \beta]$$

I need to show the other inclusion. This will be done for a suitable choice of λ. To do this, it suffices to verify that both α and β are in $\mathbb{F}[\gamma]$.

Let the minimal polynomials of α and β be $f(x)$ and $g(x)$ respectively. Let the distinct roots of $f(x)$ and $g(x)$ be $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ and $\{\beta_1, \beta_2, \cdots, \beta_m\}$ respectively. These roots are in a field which contains splitting fields of both $f(x)$ and $g(x)$. Let $\alpha = \alpha_1$ and $\beta = \beta_1$. Now define

$$h(x) \equiv f(\alpha + \lambda\beta - \lambda x) \equiv f(\gamma - \lambda x)$$

so that $h(\beta) = f(\alpha) = 0$. It follows $(x - \beta)$ divides both $h(x)$ and $g(x)$. If $(x - \eta)$ is a different linear factor of both $g(x)$ and $h(x)$ then it must be $(x - \beta_j)$ for some β_j for some $j > 1$ because these are the only factors of $g(x)$. Therefore, this would require

$$0 = h(\beta_j) = f(\alpha_1 + \lambda\beta_1 - \lambda\beta_j)$$

and so it would be the case that $\alpha_1 + \lambda\beta_1 - \lambda\beta_j = \alpha_k$ for some k. Hence

$$\lambda = \frac{\alpha_k - \alpha_1}{\beta_1 - \beta_j}$$

Now there are finitely many quotients of the above form and if λ is chosen to not be any of them, then the above cannot happen and so in this case, the only linear factor of both $g(x)$ and $h(x)$ will be $(x - \beta)$. Choose such a λ.

Let $\phi(x)$ be the minimal polynomial of β with respect to the field $\mathbb{F}[\gamma]$. Then this minimal polynomial must divide both $h(x)$ and $g(x)$ because $h(\beta) = g(\beta) = 0$. However, the only factor these two have in common is $x - \beta$ and so $\phi(x) = x - \beta$ which requires $\beta \in \mathbb{F}[\gamma]$. Now also $\alpha = \gamma - \lambda\beta$ and so $\alpha \in \mathbb{F}[\gamma]$ also. Therefore, both $\alpha, \beta \in \mathbb{F}[\gamma]$ which forces $\mathbb{F}[\alpha, \beta] \subseteq \mathbb{F}[\gamma]$. This proves the proposition in the case that $n = 2$. The general result follows right away by observing that

$$\mathbb{F}[a_1 \cdots a_n] = \mathbb{F}[a_1 \cdots a_{n-1}][a_n]$$

and using induction. ■

When you have a field $\mathbb{F}$, $\mathbb{F}(a)$ denotes the smallest field which contains both $\mathbb{F}$ and a. When a is algebraic over $\mathbb{F}$, it follows that $\mathbb{F}(a) = \mathbb{F}[a]$. The latter is easier to think about because it just involves polynomials.

7.3.4 The Lindemannn Weierstrass Theorem And Vector Spaces

As another application of the abstract concept of vector spaces, there is an amazing theorem due to Weierstrass and Lindemannn.

Theorem 7.3.34 *Suppose $a_1, \cdots, a_n$ are algebraic numbers, roots of a polynomial with rational coefficients, and suppose $\alpha_1, \cdots, \alpha_n$ are distinct algebraic numbers. Then*

$$\sum_{i=1}^{n} a_i e^{\alpha_i} \neq 0$$

In other words, the $\{e^{\alpha_1}, \cdots, e^{\alpha_n}\}$ are independent as vectors with field of scalars equal to the algebraic numbers.

There is a proof of this in the appendix. It is long and hard but only depends on elementary considerations other than some algebra involving symmetric polynomials. See Theorem F.3.5.

A number is transcendental, as opposed to algebraic, if it is not a root of a polynomial which has integer (rational) coefficients. Most numbers are this way but it is hard to verify that specific numbers are transcendental. That π is transcendental follows from

$$e^0 + e^{i\pi} = 0.$$

By the above theorem, this could not happen if π were algebraic because then $i\pi$ would also be algebraic. Recall these algebraic numbers form a field and i is clearly algebraic, being a root of $x^2 + 1$. This fact about π was first proved by Lindemannn in 1882 and then the general theorem above was proved by Weierstrass in 1885. This fact that π is transcendental solved an old problem called squaring the circle which was to construct a square with the same area as a circle using a straight edge and compass. It can be shown that the fact π is transcendental implies this problem is impossible.[1]

[1]Gilbert, the librettist of the Savoy operas, may have heard about this great achievement. In Princess Ida which opened in 1884 he has the following lines. "As for fashion they forswear it, so the say - so they say; and the circle - they will square it some fine day some fine day." Of course it had been proved impossible to do this a couple of years before.

7.4 Exercises

1. Let H denote span $\left(\begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}1\\4\\0\end{pmatrix}, \begin{pmatrix}1\\3\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}\right)$. Find the dimension of H and determine a basis.
2. Let $M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_3 = u_1 = 0\}$. Is M a subspace? Explain.
3. Let $M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_3 \geq u_1\}$. Is M a subspace? Explain.
4. Let $\mathbf{w} \in \mathbb{R}^4$ and let $M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : \mathbf{w} \cdot \mathbf{u} = 0\}$. Is M a subspace? Explain.
5. Let $M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_i \geq 0 \text{ for each } i = 1, 2, 3, 4\}$. Is M a subspace? Explain.
6. Let $\mathbf{w}, \mathbf{w}_1$ be given vectors in $\mathbb{R}^4$ and define
$$M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : \mathbf{w} \cdot \mathbf{u} = 0 \text{ and } \mathbf{w}_1 \cdot \mathbf{u} = 0\}.$$
Is M a subspace? Explain.
7. Let $M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : |u_1| \leq 4\}$. Is M a subspace? Explain.
8. Let $M = \{\mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : \sin(u_1) = 1\}$. Is M a subspace? Explain.
9. Suppose $\{\mathbf{x}_1, \cdots, \mathbf{x}_k\}$ is a set of vectors from $\mathbb{F}^n$. Show that $\mathbf{0}$ is in span $(\mathbf{x}_1, \cdots, \mathbf{x}_k)$.
10. Consider the vectors of the form
$$\left\{\begin{pmatrix}2t+3s\\s-t\\t+s\end{pmatrix} : s, t \in \mathbb{R}\right\}.$$
Is this set of vectors a subspace of $\mathbb{R}^3$? If so, explain why, give a basis for the subspace and find its dimension.
11. Consider the vectors of the form
$$\left\{\begin{pmatrix}2t+3s+u\\s-t\\t+s\\u\end{pmatrix} : s, t, u \in \mathbb{R}\right\}.$$
Is this set of vectors a subspace of $\mathbb{R}^4$? If so, explain why, give a basis for the subspace and find its dimension.
12. Consider the vectors of the form
$$\left\{\begin{pmatrix}2t+u+1\\t+3u\\t+s+v\\u\end{pmatrix} : s, t, u, v \in \mathbb{R}\right\}.$$
Is this set of vectors a subspace of $\mathbb{R}^4$? If so, explain why, give a basis for the subspace and find its dimension.
13. Let V denote the set of functions defined on $[0, 1]$. Vector addition is defined as $(f+g)(x) \equiv f(x) + g(x)$ and scalar multiplication is defined as $(\alpha f)(x) \equiv \alpha(f(x))$. Verify V is a vector space. What is its dimension, finite or infinite? Justify your answer.

14. Let V denote the set of polynomial functions defined on $[0,1]$. Vector addition is defined as $(f+g)(x) \equiv f(x) + g(x)$ and scalar multiplication is defined as $(\alpha f)(x) \equiv \alpha(f(x))$. Verify V is a vector space. What is its dimension, finite or infinite? Justify your answer.

15. Let V be the set of polynomials defined on $\mathbb{R}$ having degree no more than 4. Give a basis for this vector space.

16. Let the vectors be of the form $a + b\sqrt{2}$ where a, b are rational numbers and let the field of scalars be $\mathbb{F} = \mathbb{Q}$, the rational numbers. Show directly this is a vector space. What is its dimension? What is a basis for this vector space?

17. Let V be a vector space with field of scalars $\mathbb{F}$ and suppose $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis for V. Now let W also be a vector space with field of scalars $\mathbb{F}$. Let $L : \{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \to W$ be a function such that $L\mathbf{v}_j = \mathbf{w}_j$. Explain how L can be extended to a linear transformation mapping V to W in a unique way.

18. If you have 5 vectors in $\mathbb{F}^5$ and the vectors are linearly independent, can it always be concluded they span $\mathbb{F}^5$? Explain.

19. If you have 6 vectors in $\mathbb{F}^5$, is it possible they are linearly independent? Explain.

20. Suppose V, W are subspaces of $\mathbb{F}^n$. Show $V \cap W$ defined to be all vectors which are in both V and W is a subspace also.

21. Suppose V and W both have dimension equal to 7 and they are subspaces of a vector space of dimension 10. What are the possibilities for the dimension of $V \cap W$? **Hint:** Remember that a linear independent set can be extended to form a basis.

22. Suppose V has dimension p and W has dimension q and they are each contained in a subspace, U which has dimension equal to n where $n > \max(p, q)$. What are the possibilities for the dimension of $V \cap W$? **Hint:** Remember that a linear independent set can be extended to form a basis.

23. If $\mathbf{b} \neq \mathbf{0}$, can the solution set of $A\mathbf{x} = \mathbf{b}$ be a plane through the origin? Explain.

24. Suppose a system of equations has fewer equations than variables and you have found a solution to this system of equations. Is it possible that your solution is the only one? Explain.

25. Suppose a system of linear equations has a 2×4 augmented matrix and the last column is a pivot column. Could the system of linear equations be consistent? Explain.

26. Suppose the coefficient matrix of a system of n equations with n variables has the property that every column is a pivot column. Does it follow that the system of equations must have a solution? If so, must the solution be unique? Explain.

27. Suppose there is a unique solution to a system of linear equations. What must be true of the pivot columns in the augmented matrix.

28. State whether each of the following sets of data are possible for the matrix equation $A\mathbf{x} = \mathbf{b}$. If possible, describe the solution set. That is, tell whether there exists a unique solution no solution or infinitely many solutions.

    (a) A is a 5×6 matrix, $\operatorname{rank}(A) = 4$ and $\operatorname{rank}(A|\mathbf{b}) = 4$. **Hint:** This says $\mathbf{b}$ is in the span of four of the columns. Thus the columns are not independent.

    (b) A is a 3×4 matrix, $\operatorname{rank}(A) = 3$ and $\operatorname{rank}(A|\mathbf{b}) = 2$.

    (c) A is a 4×2 matrix, $\operatorname{rank}(A) = 4$ and $\operatorname{rank}(A|\mathbf{b}) = 4$. **Hint:** This says $\mathbf{b}$ is in the span of the columns and the columns must be independent.

    (d) A is a 5×5 matrix, $\operatorname{rank}(A) = 4$ and $\operatorname{rank}(A|\mathbf{b}) = 5$. **Hint:** This says $\mathbf{b}$ is not in the span of the columns.

    (e) A is a 4×2 matrix, $\operatorname{rank}(A) = 2$ and $\operatorname{rank}(A|\mathbf{b}) = 2$.

29. Suppose A is an $m \times n$ matrix in which $m \leq n$. Suppose also that the rank of A equals m. Show that A maps $\mathbb{F}^n$ onto $\mathbb{F}^m$. **Hint:** The vectors $\mathbf{e}_1, \cdots, \mathbf{e}_m$ occur as columns in the row reduced echelon form for A.

30. Suppose A is an $m \times n$ matrix in which $m \geq n$. Suppose also that the rank of A equals n. Show that A is one to one. **Hint:** If not, there exists a vector, $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$, and this implies at least one column of A is a linear combination of the others. Show this would require the column rank to be less than n.

31. Explain why an $n \times n$ matrix A is both one to one and onto if and only if its rank is n.

32. If you have not done this already, here it is again. It is a very important result. Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Show that

$$\dim(\ker(AB)) \leq \dim(\ker(A)) + \dim(\ker(B)).$$

Hint: Consider the subspace, $B(\mathbb{F}^p) \cap \ker(A)$ and suppose a basis for this subspace is $\{\mathbf{w}_1, \cdots, \mathbf{w}_k\}$. Now suppose $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is a basis for $\ker(B)$. Let $\{\mathbf{z}_1, \cdots, \mathbf{z}_k\}$ be such that $B\mathbf{z}_i = \mathbf{w}_i$ and argue that

$$\ker(AB) \subseteq \operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{z}_1, \cdots, \mathbf{z}_k).$$

Here is how you do this. Suppose $AB\mathbf{x} = \mathbf{0}$. Then $B\mathbf{x} \in \ker(A) \cap B(\mathbb{F}^p)$ and so $B\mathbf{x} = \sum_{i=1}^k B\mathbf{z}_i$ showing that

$$\mathbf{x} - \sum_{i=1}^k \mathbf{z}_i \in \ker(B).$$

33. Recall that every positive integer can be factored into a product of primes in a unique way. Show there must be infinitely many primes. **Hint:** Show that if you have any finite set of primes and you multiply them and then add 1, the result cannot be divisible by any of the primes in your finite set. This idea in the hint is due to Euclid who lived about 300 B.C.

34. There are lots of fields. This will give an example of a finite field. Let $\mathbb{Z}$ denote the set of integers. Thus $\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$. Also let p be a prime number. We will say that two integers, a, b are equivalent and write $a \sim b$ if $a - b$ is divisible by p. Thus they are equivalent if $a - b = px$ for some integer x. First show that $a \sim a$. Next show that if $a \sim b$ then $b \sim a$. Finally show that if $a \sim b$ and $b \sim c$ then $a \sim c$. For a an integer, denote by $[a]$ the set of all integers which is equivalent to a, the equivalence class of a. Show first that is suffices to consider only $[a]$ for $a = 0, 1, 2, \cdots, p-1$ and that for $0 \leq a < b \leq p-1$, $[a] \neq [b]$. That is, $[a] = [r]$ where $r \in \{0, 1, 2, \cdots, p-1\}$. Thus there are exactly p of these equivalence classes. **Hint:** Recall the Euclidean algorithm. For $a > 0$, $a = mp + r$ where $r < p$. Next define the following operations.

$$\begin{aligned} [a] + [b] &\equiv [a+b] \\ [a]\,[b] &\equiv [ab] \end{aligned}$$

Show these operations are well defined. That is, if $[a] = [a']$ and $[b] = [b']$, then $[a] + [b] = [a'] + [b']$ with a similar conclusion holding for multiplication. Thus for addition you need to verify $[a+b] = [a'+b']$ and for multiplication you need to verify $[ab] = [a'b']$. For example, if $p = 5$ you have $[3] = [8]$ and $[2] = [7]$. Is $[2 \times 3] = [8 \times 7]$? Is $[2+3] = [8+7]$? Clearly so in this example because when you subtract, the result is divisible by 5. So why is this so in general? Now verify that $\{[0], [1], \cdots, [p-1]\}$ with these operations is a Field. This is called the integers modulo a prime and is written $\mathbb{Z}_p$. Since there are infinitely many primes p, it follows there are infinitely many of these finite fields. **Hint:** Most of the axioms are easy once you have shown the operations are well defined. The only two which are tricky are the ones which give the existence of the additive inverse and the multiplicative inverse. Of these, the first is not hard. $-[x] = [-x]$. Since p is prime, there exist integers x, y such that $1 = px + ky$ and so $1 - ky = px$ which says $1 \sim ky$ and so $[1] = [ky]$. Now you finish the argument. What is the multiplicative identity in this collection of equivalence classes? Of course you could now consider field extensions based on these fields.

35. Suppose the field of scalars is $\mathbb{Z}_2$ described above. Show that

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

Thus the identity is a comutator. Compare this with Problem 50 on Page 187.

36. Suppose V is a vector space with field of scalars $\mathbb{F}$. Let $T \in \mathcal{L}(V, W)$, the space of linear transformations mapping V onto W where W is another vector space. Define an equivalence relation on V as follows. $\mathbf{v} \sim \mathbf{w}$ means $\mathbf{v} - \mathbf{w} \in \ker(T)$. Recall that $\ker(T) \equiv \{\mathbf{v} : T\mathbf{v} = \mathbf{0}\}$. Show this is an equivalence relation. Now for $[\mathbf{v}]$ an equivalence class define $T'[\mathbf{v}] \equiv T\mathbf{v}$. Show this is well defined. Also show that with the operations

$$\begin{aligned} [\mathbf{v}] + [\mathbf{w}] &\equiv [\mathbf{v} + \mathbf{w}] \\ \alpha\,[\mathbf{v}] &\equiv [\alpha \mathbf{v}] \end{aligned}$$

this set of equivalence classes, denoted by $V/\ker(T)$ is a vector space. Show next that $T' : V/\ker(T) \to W$ is one to one, linear, and onto. This new vector space, $V/\ker(T)$ is called a quotient space. Show its dimension equals the difference between the dimension of V and the dimension of $\ker(T)$.

37. Let V be an n dimensional vector space and let W be a subspace. Generalize the above problem to define and give properties of V/W. What is its dimension? What is a basis?

38. If $\mathbb{F}$ and $\mathbb{G}$ are two fields and $\mathbb{F} \subseteq \mathbb{G}$, can you consider $\mathbb{G}$ as a vector space with field of scalars $\mathbb{F}$? Explain.

39. Let $\mathbb{A}$ denote the real roots of polynomials in $\mathbb{Q}[x]$. Show $\mathbb{A}$ can be considered a vector space with field of scalars $\mathbb{Q}$. What is the dimension of this vector space, finite or infinite?

40. As mentioned, for distinct algebraic numbers α_i, the complex numbers $\{e^{\alpha_i}\}_{i=1}^n$ are linearly independent over the field of scalars $\mathbb{A}$ where $\mathbb{A}$ denotes the algebraic numbers, those which are roots of a polynomial having integer (rational) coefficients. What is the dimension of the vector space $\mathbb{C}$ with field of scalars $\mathbb{A}$, finite or infinite? If the field of scalars were $\mathbb{C}$ instead of $\mathbb{A}$, would this change? What if the field of scalars were $\mathbb{R}$?

41. Suppose $\mathbb{F}$ is a countable field and let $\mathbb{A}$ be the algebraic numbers, those numbers in $\mathbb{G}$ which are roots of a polynomial in $\mathbb{F}[x]$. Show $\mathbb{A}$ is also countable.

42. This problem is on partial fractions. Suppose you have

$$R(x) = \frac{p(x)}{q_1(x) \cdots q_m(x)}, \text{ degree of } p(x) < \text{ degree of denominator.}$$

where the polynomials $q_i(x)$ are relatively prime and all the polynomials $p(x)$ and $q_i(x)$ have coefficients in a field of scalars $\mathbb{F}$. Thus there exist polynomials $a_i(x)$ having coefficients in $\mathbb{F}$ such that

$$1 = \sum_{i=1}^{m} a_i(x) q_i(x)$$

Explain why

$$R(x) = \frac{p(x) \sum_{i=1}^{m} a_i(x) q_i(x)}{q_1(x) \cdots q_m(x)} = \sum_{i=1}^{m} \frac{a_i(x) p(x)}{\prod_{j \neq i} q_j(x)}$$

Now continue doing this on each term in the above sum till finally you obtain an expression of the form

$$\sum_{i=1}^{m} \frac{b_i(x)}{q_i(x)}$$

Using the Euclidean algorithm for polynomials, explain why the above is of the form

$$M(x) + \sum_{i=1}^{m} \frac{r_i(x)}{q_i(x)}$$

where the degree of each $r_i(x)$ is less than the degree of $q_i(x)$ and $M(x)$ is a polynomial. Now argue that $M(x) = 0$. From this explain why the usual partial fractions expansion of calculus must be true. You can use the fact that every polynomial having real coefficients factors into a product of irreducible quadratic polynomials and linear polynomials having real coefficients. This follows from the fundamental theorem of algebra in the appendix.

43. Suppose $\{f_1, \cdots, f_n\}$ is an independent set of smooth functions defined on some interval (a, b). Now let A be an invertible $n \times n$ matrix. Define new functions $\{g_1, \cdots, g_n\}$ as follows.

$$\begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} = A \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

Is it the case that $\{g_1, \cdots, g_n\}$ is also independent? Explain why.

44. A number is transcendental if it is not the root of any nonzero polynomial with rational coefficients. As mentioned, there are many known transcendental numbers. Suppose α is a real transcendental number. Show that $\{1, \alpha, \alpha^2, \cdots\}$ is a linearly independent set of real numbers if the field of scalars is the rational numbers.

Chapter 8

Linear Transformations

8.1 Matrix Multiplication As A Linear Transformation

Definition 8.1.1 *Let V and W be two finite dimensional vector spaces. A function, L which maps V to W is called a linear transformation and written $L \in \mathcal{L}(V,W)$ if for all scalars α and β, and vectors v,w,*

$$L(\alpha v+\beta w) = \alpha L(v) + \beta L(w).$$

An example of a linear transformation is familiar matrix multiplication. Let $A = (a_{ij})$ be an $m \times n$ matrix. Then an example of a linear transformation $L : \mathbb{F}^n \to \mathbb{F}^m$ is given by

$$(L\mathbf{v})_i \equiv \sum_{j=1}^{n} a_{ij} v_j.$$

Here

$$\mathbf{v} \equiv \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{F}^n.$$

8.2 $\mathcal{L}(V,W)$ As A Vector Space

Definition 8.2.1 *Given $L, M \in \mathcal{L}(V,W)$ define a new element of $\mathcal{L}(V,W)$, denoted by $L+M$ according to the rule*[1]

$$(L+M)\,v \equiv Lv + Mv.$$

For α a scalar and $L \in \mathcal{L}(V,W)$, define $\alpha L \in \mathcal{L}(V,W)$ by

$$\alpha L(v) \equiv \alpha(Lv).$$

You should verify that all the axioms of a vector space hold for $\mathcal{L}(V,W)$ with the above definitions of vector addition and scalar multiplication. What about the dimension of $\mathcal{L}(V,W)$?

Before answering this question, here is a useful lemma. It gives a way to define linear transformations and a way to tell when two of them are equal.

Lemma 8.2.2 *Let V and W be vector spaces and suppose $\{v_1, \cdots, v_n\}$ is a basis for V. Then if $L : V \to W$ is given by $Lv_k = w_k \in W$ and*

$$L\left(\sum_{k=1}^{n} a_k v_k\right) \equiv \sum_{k=1}^{n} a_k L v_k = \sum_{k=1}^{n} a_k w_k$$

then L is well defined and is in $\mathcal{L}(V,W)$. Also, if L, M are two linear transformations such that $Lv_k = Mv_k$ for all k, then $M = L$.

Proof: L is well defined on V because, since $\{v_1, \cdots, v_n\}$ is a basis, there is exactly one way to write a given vector of V as a linear combination. Next, observe that L is obviously linear from the definition. If L, M are equal on the basis, then if $\sum_{k=1}^{n} a_k v_k$ is an arbitrary vector of V,

$$L\left(\sum_{k=1}^{n} a_k v_k\right) = \sum_{k=1}^{n} a_k L v_k = \sum_{k=1}^{n} a_k M v_k = M\left(\sum_{k=1}^{n} a_k v_k\right)$$

and so $L = M$ because they give the same result for every vector in V. ■

The message is that when you define a linear transformation, it suffices to tell what it does to a basis.

Theorem 8.2.3 *Let V and W be finite dimensional linear spaces of dimension n and m respectively Then* $\dim(\mathcal{L}(V,W)) = mn$.

Proof: Let two sets of bases be

$$\{v_1, \cdots, v_n\} \text{ and } \{w_1, \cdots, w_m\}$$

for V and W respectively. Using Lemma 8.2.2, let $w_i v_j \in \mathcal{L}(V,W)$ be the linear transformation defined on the basis, $\{v_1, \cdots, v_n\}$, by

$$w_i v_k(v_j) \equiv w_i \delta_{jk}$$

where $\delta_{ik} = 1$ if $i = k$ and 0 if $i \neq k$. I will show that $L \in \mathcal{L}(V,W)$ is a linear combination of these special linear transformations called dyadics.

Then let $L \in \mathcal{L}(V,W)$. Since $\{w_1, \cdots, w_m\}$ is a basis, there exist constants, d_{jk} such that

$$Lv_r = \sum_{j=1}^{m} d_{jr} w_j$$

[1] Note that this is the standard way of defining the sum of two functions.

Now consider the following sum of dyadics.

$$\sum_{j=1}^{m}\sum_{i=1}^{n} d_{ji} w_j v_i$$

Apply this to v_r. This yields

$$\sum_{j=1}^{m}\sum_{i=1}^{n} d_{ji} w_j v_i (v_r) = \sum_{j=1}^{m}\sum_{i=1}^{n} d_{ji} w_j \delta_{ir} = \sum_{j=1}^{m} d_{jr} w_i = L v_r$$

Therefore, $L = \sum_{j=1}^{m}\sum_{i=1}^{n} d_{ji} w_j v_i$ showing the span of the dyadics is all of $\mathcal{L}(V,W)$.

Now consider whether these dyadics form a linearly independent set. Suppose

$$\sum_{i,k} d_{ik} w_i v_k = \mathbf{0}.$$

Are all the scalars d_{ik} equal to 0?

$$\mathbf{0} = \sum_{i,k} d_{ik} w_i v_k (v_l) = \sum_{i=1}^{m} d_{il} w_i$$

and so, since $\{w_1, \cdots, w_m\}$ is a basis, $d_{il} = 0$ for each $i = 1, \cdots, m$. Since l is arbitrary, this shows $d_{il} = 0$ for all i and l. Thus these linear transformations form a basis and this shows that the dimension of $\mathcal{L}(V,W)$ is mn as claimed because there are m choices for the w_i and n choices for the v_j. ■

8.3 The Matrix Of A Linear Transformation

Definition 8.3.1 *In Theorem 8.2.3, the matrix of the linear transformation* $L \in \mathcal{L}(V,W)$ *with respect to the ordered bases* $\beta \equiv \{v_1, \cdots, v_n\}$ *for* V *and* $\gamma \equiv \{w_1, \cdots, w_m\}$ *for* W *is defined to be* $[L]$ *where* $[L]_{ij} = d_{ij}$. *Thus this matrix is defined by* $L = \sum_{i,j} [L]_{ij} w_i v_i$. *When it is desired to feature the bases* β, γ, *this matrix will be denoted as* $[L]_{\gamma\beta}$. *When there is only one basis* β, *this is denoted as* $[L]_\beta$.

If V is an n dimensional vector space and $\beta = \{v_1, \cdots, v_n\}$ is a basis for V, there exists a linear map

$$q_\beta : \mathbb{F}^n \to V$$

defined as

$$q_\beta(\mathbf{a}) \equiv \sum_{i=1}^{n} a_i v_i$$

where

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i=1}^{n} a_i \mathbf{e}_i,$$

for $\mathbf{e}_i$ the standard basis vectors for $\mathbb{F}^n$ consisting of $\begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}^T$. Thus the 1 is in the i^{th} position and the other entries are 0. Conversely, if $q : \mathbb{F}^n \to V$ is one to one, onto, and linear, it must be of the form just described. Just let $v_i \equiv q(\mathbf{e}_i)$.

It is clear that q defined in this way, is one to one, onto, and linear. For $v \in V$, $q_\beta^{-1}(v)$ is a vector in $\mathbb{F}^n$ called the component vector of v with respect to the basis $\{v_1, \cdots, v_n\}$.

Proposition 8.3.2 *The matrix of a linear transformation with respect to ordered bases* β, γ *as described above is characterized by the requirement that multiplication of the components of* v *by* $[L]_{\gamma\beta}$ *gives the components of* Lv.

Proof: This happens because by definition, if $v = \sum_i x_i v_i$, then

$$Lv = \sum_i x_i L v_i \equiv \sum_i \sum_j [L]_{ji} x_i w_j = \sum_j \sum_i [L]_{ji} x_i w_j$$

and so the j^{th} component of Lv is $\sum_i [L]_{ji} x_i$, the j^{th} component of the matrix times the component vector of v. Could there be some other matrix which will do this? No, because if such a matrix is M, then for any $\mathbf{x}$, it follows from what was just shown that $[L]\mathbf{x} = M\mathbf{x}$. Hence $[L] = M$. ■

The above proposition shows that the following diagram determines the matrix of a linear transformation. Here q_β and q_γ are the maps defined above with reference to the ordered bases, $\{v_1, \cdots, v_n\}$ and $\{w_1, \cdots, w_m\}$ respectively.

$$\begin{array}{ccccc} & & L & & \\ \beta = \{v_1, \cdots, v_n\} & V & \rightarrow & W & \{w_1, \cdots, w_m\} = \gamma \\ & q_\beta \uparrow & \circ & \uparrow q_\gamma & \\ & \mathbb{F}^n & \rightarrow & \mathbb{F}^m & \\ & & [L]_{\gamma\beta} & & \end{array} \tag{8.1}$$

In terms of this diagram, the matrix $[L]_{\gamma\beta}$ is the matrix chosen to make the diagram "commute". It may help to write the description of $[L]_{\gamma\beta}$ in the form

$$\begin{pmatrix} Lv_1 & \cdots & Lv_n \end{pmatrix} = \begin{pmatrix} w_1 & \cdots & w_m \end{pmatrix} [L]_{\gamma\beta} \tag{8.2}$$

with the understanding that you do the multiplications in a formal manner just as you would if everything were numbers. If this helps, use it. If it does not help, ignore it.

Example 8.3.3 *Let*

$$V \equiv \{ \text{ polynomials of degree 3 or less} \},$$

$$W \equiv \{ \text{ polynomials of degree 2 or less} \},$$

and $L \equiv D$ where D is the differentiation operator. A basis for V is $\beta = \{1, x, x^2, x^3\}$ and a basis for W is $\gamma = \{1, x, x^2\}$.

What is the matrix of this linear transformation with respect to this basis? Using 8.2,

$$\begin{pmatrix} 0 & 1 & 2x & 3x^2 \end{pmatrix} = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} [D]_{\gamma\beta}.$$

It follows from this that the first column of $[D]_{\gamma\beta}$ is

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The next three columns of $[D]_{\gamma\beta}$ are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$

and so

$$[D]_{\gamma\beta} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Now consider the important case where $V = \mathbb{F}^n$, $W = \mathbb{F}^m$, and the basis chosen is the standard basis of vectors $\mathbf{e}_i$ described above.

$$\beta = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\},\ \gamma = \{\mathbf{e}_1, \cdots, \mathbf{e}_m\}$$

Let L be a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$ and let A be the matrix of the transformation with respect to these bases. In this case the coordinate maps q_β and q_γ are simply the identity maps on $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively, and can be accomplished by simply multiplying by the appropriate sized identity matrix. The requirement that A is the matrix of the transformation amounts to

$$L\mathbf{b} = A\mathbf{b}$$

What about the situation where different pairs of bases are chosen for V and W? How are the two matrices with respect to these choices related? Consider the following diagram which illustrates the situation.

$$\begin{array}{ccc} \mathbb{F}^n & \underrightarrow{A_2} & \mathbb{F}^m \\ q_{\beta_2} \downarrow & \circ & q_{\gamma_2} \downarrow \\ V & \underrightarrow{L} & W \\ q_{\beta_1} \uparrow & \circ & q_{\gamma_1} \uparrow \\ \mathbb{F}^n & \underrightarrow{A_1} & \mathbb{F}^m \end{array}$$

In this diagram q_{β_i} and q_{γ_i} are coordinate maps as described above. From the diagram,

$$q_{\gamma_1}^{-1} q_{\gamma_2} A_2 q_{\beta_2}^{-1} q_{\beta_1} = A_1,$$

where $q_{\beta_2}^{-1} q_{\beta_1}$ and $q_{\gamma_1}^{-1} q_{\gamma_2}$ are one to one, onto, and linear maps which may be accomplished by multiplication by a square matrix. Thus there exist matrices P, Q such that $P : \mathbb{F}^n \to \mathbb{F}^n$ and $Q : \mathbb{F}^m \to \mathbb{F}^m$ are invertible and

$$PA_2Q = A_1.$$

Example 8.3.4 *Let $\beta \equiv \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ and $\gamma \equiv \{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ be two bases for V. Let L be the linear transformation which maps $\mathbf{v}_i$ to $\mathbf{w}_i$. Find $[L]_{\gamma\beta}$. In case $V = \mathbb{F}^n$ and letting $\delta = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, the usual basis for $\mathbb{F}^n$, find $[L]_\delta$.*

Letting δ_{ij} be the symbol which equals 1 if $i = j$ and 0 if $i \neq j$, it follows that $L = \sum_{i,j} \delta_{ij} \mathbf{w}_i \mathbf{v}_j$ and so $[L]_{\gamma\beta} = I$ the identity matrix. For the second part, you must have

$$\begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_n \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} [L]_\delta$$

and so

$$[L]_\delta = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_n \end{pmatrix}$$

where $\begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_n \end{pmatrix}$ is the $n \times n$ matrix having i^{th} column equal to $\mathbf{w}_i$.

Definition 8.3.5 *In the special case where $V = W$ and only one basis is used for $V = W$, this becomes*

$$q_{\beta_1}^{-1} q_{\beta_2} A_2 q_{\beta_2}^{-1} q_{\beta_1} = A_1.$$

Letting S be the matrix of the linear transformation $q_{\beta_2}^{-1} q_{\beta_1}$ with respect to the standard basis vectors in $\mathbb{F}^n$,

$$S^{-1} A_2 S = A_1. \tag{8.3}$$

When this occurs, A_1 is said to be similar to A_2 and $A \to S^{-1}AS$ is called a similarity transformation.

Recall the following.

Definition 8.3.6 *Let S be a set. The symbol $\sim$ is called an equivalence relation on S if it satisfies the following axioms.*

1. *$x \sim x$ for all $x \in S$. (Reflexive)*
2. *If $x \sim y$ then $y \sim x$. (Symmetric)*
3. *If $x \sim y$ and $y \sim z$, then $x \sim z$. (Transitive)*

Definition 8.3.7 *$[x]$ denotes the set of all elements of S which are equivalent to x and $[x]$ is called the equivalence class determined by x or just the equivalence class of x.*

Also recall the notion of equivalence classes.

Theorem 8.3.8 *Let $\sim$ be an equivalence class defined on a set S and let $\mathcal{H}$ denote the set of equivalence classes. Then if $[x]$ and $[y]$ are two of these equivalence classes, either $x \sim y$ and $[x] = [y]$ or it is not true that $x \sim y$ and $[x] \cap [y] = \emptyset$.*

Theorem 8.3.9 *In the vector space of $n \times n$ matrices, define*

$$A \sim B$$

if there exists an invertible matrix S such that

$$A = S^{-1}BS.$$

Then $\sim$ is an equivalence relation and $A \sim B$ if and only if whenever V is an n dimensional vector space, there exists $L \in \mathcal{L}(V, V)$ and bases $\{v_1, \cdots, v_n\}$ and $\{w_1, \cdots, w_n\}$ such that A is the matrix of L with respect to $\{v_1, \cdots, v_n\}$ and B is the matrix of L with respect to $\{w_1, \cdots, w_n\}$.

Proof: $A \sim A$ because $S = I$ works in the definition. If $A \sim B$, then $B \sim A$, because

$$A = S^{-1}BS$$

implies $B = SAS^{-1}$. If $A \sim B$ and $B \sim C$, then

$$A = S^{-1}BS, \ B = T^{-1}CT$$

and so

$$A = S^{-1}T^{-1}CTS = (TS)^{-1}\, CTS$$

which implies $A \sim C$. This verifies the first part of the conclusion.

Now let V be an n dimensional vector space, $A \sim B$ so $A = S^{-1}BS$ and pick a basis for V,

$$\beta \equiv \{v_1, \cdots, v_n\}.$$

Define $L \in \mathcal{L}(V, V)$ by

$$Lv_i \equiv \sum_j a_{ji} v_j$$

where $A = (a_{ij})$. Thus A is the matrix of the linear transformation L. Consider the diagram

$$\begin{array}{ccc} \mathbb{F}^n & \underrightarrow{B} & \mathbb{F}^n \\ q_\gamma \downarrow & \circ & q_\gamma \downarrow \\ V & \underrightarrow{L} & V \\ q_\beta \uparrow & \circ & q_\beta \uparrow \\ \mathbb{F}^n & \underrightarrow{A} & \mathbb{F}^n \end{array}$$

where q_γ is chosen to make the diagram commute. Thus we need $S = q_\gamma^{-1} q_\beta$ which requires

$$q_\gamma = q_\beta S^{-1}$$

Then it follows that B is the matrix of L with respect to the basis

$$\{q_\gamma \mathbf{e}_1, \cdots, q_\gamma \mathbf{e}_n\} \equiv \{w_1, \cdots, w_n\}.$$

That is, A and B are matrices of the same linear transformation L. Conversely, suppose whenever V is an n dimensional vector space, there exists $L \in \mathcal{L}(V, V)$ and bases $\{v_1, \cdots, v_n\}$ and $\{w_1, \cdots, w_n\}$ such that A is the matrix of L with respect to $\{v_1, \cdots, v_n\}$ and B is the matrix of L with respect to $\{w_1, \cdots, w_n\}$. Then it was shown above that $A \sim B$. ■

What if the linear transformation consists of multiplication by a matrix A and you want to find the matrix of this linear transformation with respect to another basis? Is there an easy way to do it? The next proposition considers this.

Proposition 8.3.10 *Let A be an $m \times n$ matrix and let L be the linear transformation which is defined by*

$$L\left(\sum_{k=1}^{n} x_k \mathbf{e}_k\right) \equiv \sum_{k=1}^{n} (A\mathbf{e}_k) x_k \equiv \sum_{i=1}^{m}\sum_{k=1}^{n} A_{ik} x_k \mathbf{e}_i$$

In simple language, to find $L\mathbf{x}$, you multiply on the left of $\mathbf{x}$ by A. (A is the matrix of L with respect to the standard basis.) Then the matrix M of this linear transformation with respect to the bases $\beta = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ for $\mathbb{F}^n$ and $\gamma = \{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ for $\mathbb{F}^m$ is given by

$$M = \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_m \end{pmatrix}^{-1} A \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}$$

where $\begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_m \end{pmatrix}$ is the $m \times m$ matrix which has $\mathbf{w}_j$ as its j^{th} column.

Proof: Consider the following diagram.

$$\begin{array}{ccc} & L & \\ \mathbb{F}^n & \to & \mathbb{F}^m \\ q_\beta \uparrow & \circ & \uparrow q_\gamma \\ \mathbb{F}^n & \to & \mathbb{F}^m \\ & M & \end{array}$$

Here the coordinate maps are defined in the usual way. Thus

$$q_\beta \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T \equiv \sum_{i=1}^{n} x_i \mathbf{u}_i.$$

Therefore, q_β can be considered the same as multiplication of a vector in $\mathbb{F}^n$ on the left by the matrix $\begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}$. Similar considerations apply to q_γ. Thus it is desired to have the following for an arbitrary $\mathbf{x} \in \mathbb{F}^n$.

$$A \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_n \end{pmatrix} M\mathbf{x}$$

Therefore, the conclusion of the proposition follows. ■

In the special case where $m = n$ and $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$ and $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ is an orthonormal basis and you want M, the matrix of L with respect to this new orthonormal basis, it follows from the above that

$$M = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_m \end{pmatrix}^* A \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} = U^* A U$$

where U is a unitary matrix. Thus matrices with respect to two orthonormal bases are unitarily similar.

Definition 8.3.11 *An $n \times n$ matrix A, is diagonalizable if there exists an invertible $n \times n$ matrix S such that $S^{-1}AS = D$, where D is a diagonal matrix. Thus D has zero entries everywhere except on the main diagonal. Write $\operatorname{diag}(\lambda_1 \cdots, \lambda_n)$ to denote the diagonal matrix having the λ_i down the main diagonal.*

The following theorem is of great significance.

Theorem 8.3.12 *Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if $\mathbb{F}^n$ has a basis of eigenvectors of A. In this case, S of Definition 8.3.11 consists of the $n \times n$ matrix whose columns are the eigenvectors of A and $D = \text{diag}(\lambda_1, \cdots, \lambda_n)$.*

Proof: Suppose first that $\mathbb{F}^n$ has a basis of eigenvectors, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ where $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$. Then let S denote the matrix $\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$ and let $S^{-1} \equiv \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix}$ where

$$\mathbf{u}_i^T \mathbf{v}_j = \delta_{ij} \equiv \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}.$$

S^{-1} exists because S has rank n. Then from block multiplication,

$$S^{-1}AS = \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} (A\mathbf{v}_1 \cdots A\mathbf{v}_n) = \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} (\lambda_1 \mathbf{v}_1 \cdots \lambda_n \mathbf{v}_n)$$

$$= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = D.$$

Next suppose A is diagonalizable so $S^{-1}AS = D \equiv \text{diag}(\lambda_1, \cdots, \lambda_n)$. Then the columns of S form a basis because S^{-1} is given to exist. It only remains to verify that these columns of S are eigenvectors. But letting $S = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$, $AS = SD$ and so $\begin{pmatrix} A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{v}_1 & \cdots & \lambda_n \mathbf{v}_n \end{pmatrix}$ which shows that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$. ■

It makes sense to speak of the determinant of a linear transformation as described in the following corollary.

Corollary 8.3.13 *Let $L \in \mathcal{L}(V, V)$ where V is an n dimensional vector space and let A be the matrix of this linear transformation with respect to a basis on V. Then it is possible to define*

$$\det(L) \equiv \det(A).$$

Proof: Each choice of basis for V determines a matrix for L with respect to the basis. If A and B are two such matrices, it follows from Theorem 8.3.9 that

$$A = S^{-1}BS$$

and so

$$\det(A) = \det\left(S^{-1}\right)\det(B)\det(S).$$

But

$$1 = \det(I) = \det\left(S^{-1}S\right) = \det(S)\det\left(S^{-1}\right)$$

and so

$$\det(A) = \det(B) \quad ■$$

Definition 8.3.14 *Let $A \in \mathcal{L}(X, Y)$ where X and Y are finite dimensional vector spaces. Define* $\text{rank}(A)$ *to equal the dimension of $A(X)$.*

The following theorem explains how the rank of A is related to the rank of the matrix of A.

Theorem 8.3.15 *Let $A \in \mathcal{L}(X, Y)$. Then* $\text{rank}(A) = \text{rank}(M)$ *where M is the matrix of A taken with respect to a pair of bases for the vector spaces X, and Y.*

Proof: Recall the diagram which describes what is meant by the matrix of A. Here the two bases are as indicated.

$$\begin{array}{ccccc} \beta = \{v_1, \cdots, v_n\} & X & \underrightarrow{A} & Y & \{w_1, \cdots, w_m\} = \gamma \\ & q_\beta \uparrow & \circ & \uparrow q_\gamma & \\ & \mathbb{F}^n & \underrightarrow{M} & \mathbb{F}^m & \end{array}$$

Let $\{Ax_1, \cdots, Ax_r\}$ be a basis for AX. Thus

$$\left\{q_\gamma M q_\beta^{-1} x_1, \cdots, q_\gamma M q_\beta^{-1} x_r\right\}$$

is a basis for AX. It follows that

$$\left\{M q_X^{-1} x_1, \cdots, M q_X^{-1} x_r\right\}$$

is linearly independent and so $\text{rank}\,(A) \leq \text{rank}\,(M)$. However, one could interchange the roles of M and A in the above argument and thereby turn the inequality around. ■

The following result is a summary of many concepts.

Theorem 8.3.16 *Let $L \in \mathcal{L}(V, V)$ where V is a finite dimensional vector space. Then the following are equivalent.*

1. *L is one to one.*
2. *L maps a basis to a basis.*
3. *L is onto.*
4. $\det(L) \neq 0$
5. *If $Lv = 0$ then $v = 0$.*

Proof: Suppose first L is one to one and let $\beta = \{v_i\}_{i=1}^n$ be a basis. Then if $\sum_{i=1}^n c_i L v_i = 0$ it follows $L\left(\sum_{i=1}^n c_i v_i\right) = 0$ which means that since $L(0) = 0$, and L is one to one, it must be the case that $\sum_{i=1}^n c_i v_i = 0$. Since $\{v_i\}$ is a basis, each $c_i = 0$ which shows $\{Lv_i\}$ is a linearly independent set. Since there are n of these, it must be that this is a basis.

Now suppose 2.). Then letting $\{v_i\}$ be a basis, and $y \in V$, it follows from part 2.) that there are constants, $\{c_i\}$ such that $y = \sum_{i=1}^n c_i L v_i = L\left(\sum_{i=1}^n c_i v_i\right)$. Thus L is onto. It has been shown that 2.) implies 3.).

Now suppose 3.). Then the operation consisting of multiplication by the matrix of L, $[L]$, must be onto. However, the vectors in $\mathbb{F}^n$ so obtained, consist of linear combinations of the columns of $[L]$. Therefore, the column rank of $[L]$ is n. By Theorem 3.3.23 this equals the determinant rank and so $\det([L]) \equiv \det(L) \neq 0$.

Now assume 4.) If $Lv = 0$ for some $v \neq 0$, it follows that $[L]\,\mathbf{x} = 0$ for some $\mathbf{x} \neq \mathbf{0}$. Therefore, the columns of $[L]$ are linearly dependent and so by Theorem 3.3.23, $\det([L]) = \det(L) = 0$ contrary to 4.). Therefore, 4.) implies 5.).

Now suppose 5.) and suppose $Lv = Lw$. Then $L(v - w) = 0$ and so by 5.), $v - w = 0$ showing that L is one to one. ■

Also it is important to note that composition of linear transformations corresponds to multiplication of the matrices. Consider the following diagram in which $[A]_{\gamma\beta}$ denotes the matrix of A relative to the bases γ on Y and β on X, $[B]_{\delta\gamma}$ defined similarly.

$$\begin{array}{ccccc}
X & \underrightarrow{A} & Y & \underrightarrow{B} & Z \\
q_\beta \uparrow & \circ & \uparrow q_\gamma & \circ & \uparrow q_\delta \\
\mathbb{F}^n & \underrightarrow{[A]_{\gamma\beta}} & \mathbb{F}^m & \underrightarrow{[B]_{\delta\gamma}} & \mathbb{F}^p
\end{array}$$

where A and B are two linear transformations, $A \in \mathcal{L}(X, Y)$ and $B \in \mathcal{L}(Y, Z)$. Then $B \circ A \in \mathcal{L}(X, Z)$ and so it has a matrix with respect to bases given on X and Z, the coordinate maps for these bases being q_β and q_δ respectively. Then

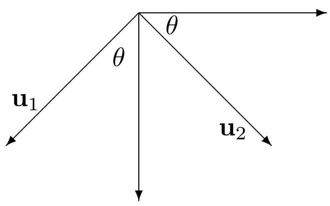
$$B \circ A = q_\delta\, [B]_{\delta\gamma}\, q_\gamma^{-1} q_\gamma\, [A]_{\gamma\beta}\, q_\beta^{-1} = q_\delta\, [B]_{\delta\gamma}\, [A]_{\gamma\beta}\, q_\beta^{-1}.$$

But this shows that $[B]_{\delta\gamma}\, [A]_{\gamma\beta}$ plays the role of $[B \circ A]_{\delta\beta}$, the matrix of $B \circ A$. Hence the matrix of $B \circ A$ equals the product of the two matrices $[A]_{\gamma\beta}$ and $[B]_{\delta\gamma}$. Of course it is interesting to note that although $[B \circ A]_{\delta\beta}$ must be unique, the matrices, $[A]_{\gamma\beta}$ and $[B]_{\delta\gamma}$ are not unique because they depend on γ, the basis chosen for Y.

Theorem 8.3.17 *The matrix of the composition of linear transformations equals the product of the matrices of these linear transformations.*

8.3.1 Rotations About A Given Vector

As an application, I will consider the problem of rotating counter clockwise about a given unit vector which is possibly not one of the unit vectors in coordinate directions. First consider a pair of perpendicular unit vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$ and the problem of rotating in the counterclockwise direction about $\mathbf{u}_3$ where $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$ so that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ forms a right handed orthogonal coordinate system. Thus the vector $\mathbf{u_3}$ is coming out of the page.



Let T denote the desired rotation. Then

$$T(a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3) = aT\mathbf{u}_1 + bT\mathbf{u}_2 + cT\mathbf{u}_3$$

$$= (a\cos\theta - b\sin\theta)\,\mathbf{u}_1 + (a\sin\theta + b\cos\theta)\,\mathbf{u}_2 + c\mathbf{u}_3.$$

Thus in terms of the basis $\gamma \equiv \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, the matrix of this transformation is

$$[T]_\gamma \equiv \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

I want to obtain the matrix of the transformation in terms of the usual basis $\beta \equiv \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ because it is in terms of this basis that we usually deal with vectors. From Proposition 8.3.10, if $[T]_\beta$ is this matrix,

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}^{-1} [T]_\beta \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}$$

and so you can solve for $[T]_\beta$ if you know the $\mathbf{u}_i$.

Recall why this is so.

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{[T]_\gamma} & \mathbb{R}^3 \\ q_\gamma \downarrow & \circ & q_\gamma \downarrow \\ \mathbb{R}^3 & \xrightarrow{T} & \mathbb{R}^3 \\ I \uparrow & \circ & I \uparrow \\ \mathbb{R}^3 & \xrightarrow{[T]_\beta} & \mathbb{R}^3 \end{array}$$

The map q_γ is accomplished by a multiplication on the left by $\begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}$. Thus

$$[T]_\beta = q_\gamma [T]_\gamma q_\gamma^{-1} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix} [T]_\gamma \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}^{-1}.$$

Suppose the unit vector $\mathbf{u}_3$ about which the counterclockwise rotation takes place is (a, b, c). Then I obtain vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$ such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a right handed orthonormal system with $\mathbf{u}_3 = (a, b, c)$ and then use the above result. It is of course somewhat arbitrary how this is accomplished. I will assume however, that $|c| \neq 1$ since otherwise you are looking at either clockwise or counter clockwise rotation about the positive z axis and this is a problem which has been dealt with earlier. (If $c = -1$, it amounts to clockwise rotation about the positive z axis while if $c = 1$, it is counter clockwise rotation about the positive z axis.)

Then let $\mathbf{u}_3 = (a, b, c)$ and $\mathbf{u}_2 \equiv \frac{1}{\sqrt{a^2+b^2}}(b, -a, 0)$. This one is perpendicular to $\mathbf{u}_3$. If $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is to be a right hand system it is necessary to have

$$\mathbf{u}_1 = \mathbf{u}_2 \times \mathbf{u}_3 = \frac{1}{\sqrt{(a^2+b^2)(a^2+b^2+c^2)}}\left(-ac, -bc, a^2+b^2\right)$$

Now recall that $\mathbf{u}_3$ is a unit vector and so the above equals

$$\frac{1}{\sqrt{(a^2+b^2)}}\left(-ac, -bc, a^2+b^2\right)$$

Then from the above, A is given by

$$\begin{pmatrix} \frac{-ac}{\sqrt{(a^2+b^2)}} & \frac{b}{\sqrt{a^2+b^2}} & a \\ \frac{-bc}{\sqrt{(a^2+b^2)}} & \frac{-a}{\sqrt{a^2+b^2}} & b \\ \sqrt{a^2+b^2} & 0 & c \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{-ac}{\sqrt{(a^2+b^2)}} & \frac{b}{\sqrt{a^2+b^2}} & a \\ \frac{-bc}{\sqrt{(a^2+b^2)}} & \frac{-a}{\sqrt{a^2+b^2}} & b \\ \sqrt{a^2+b^2} & 0 & c \end{pmatrix}^{-1}$$

Of course the matrix is an orthogonal matrix so it is easy to take the inverse by simply taking the transpose. Then doing the computation and then some simplification yields

$$= \begin{pmatrix} a^2 + \left(1 - a^2\right)\cos\theta & ab\left(1 - \cos\theta\right) - c\sin\theta & ac\left(1 - \cos\theta\right) + b\sin\theta \\ ab\left(1 - \cos\theta\right) + c\sin\theta & b^2 + \left(1 - b^2\right)\cos\theta & bc\left(1 - \cos\theta\right) - a\sin\theta \\ ac\left(1 - \cos\theta\right) - b\sin\theta & bc\left(1 - \cos\theta\right) + a\sin\theta & c^2 + \left(1 - c^2\right)\cos\theta \end{pmatrix}. \quad (8.4)$$

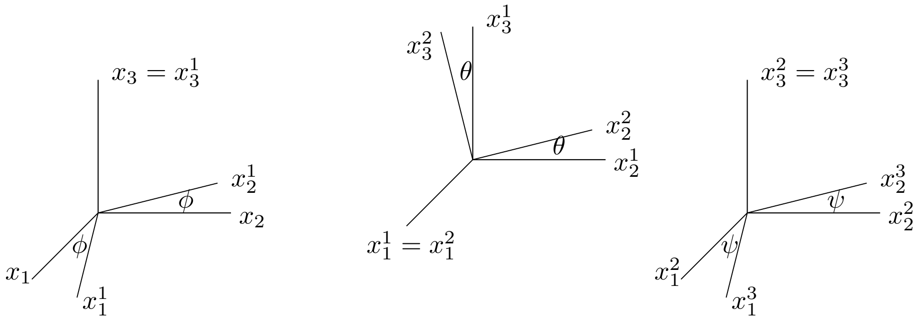
With this, it is clear how to rotate clockwise about the unit vector, (a, b, c). Just rotate counter clockwise through an angle of $-\theta$. Thus the matrix for this clockwise rotation is just

$$= \begin{pmatrix} a^2 + \left(1 - a^2\right)\cos\theta & ab\left(1 - \cos\theta\right) + c\sin\theta & ac\left(1 - \cos\theta\right) - b\sin\theta \\ ab\left(1 - \cos\theta\right) - c\sin\theta & b^2 + \left(1 - b^2\right)\cos\theta & bc\left(1 - \cos\theta\right) + a\sin\theta \\ ac\left(1 - \cos\theta\right) + b\sin\theta & bc\left(1 - \cos\theta\right) - a\sin\theta & c^2 + \left(1 - c^2\right)\cos\theta \end{pmatrix}.$$

In deriving 8.4 it was assumed that $c \neq \pm 1$ but even in this case, it gives the correct answer. Suppose for example that $c = 1$ so you are rotating in the counter clockwise direction about the positive z axis. Then a, b are both equal to zero and 8.4 reduces to 2.24.

8.3.2 The Euler Angles

An important application of the above theory is to the Euler angles, important in the mechanics of rotating bodies. Lagrange studied these things back in the 1700's. To describe the Euler angles consider the following picture in which x_1, x_2 and x_3 are the usual coordinate axes fixed in space and the axes labeled with a superscript denote other coordinate axes. Here is the picture.



We obtain ϕ by rotating counter clockwise about the fixed x_3 axis. Thus this rotation has the matrix

$$\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv M_1(\phi)$$

Next rotate counter clockwise about the x_1^1 axis which results from the first rotation through an angle of θ. Thus it is desired to rotate counter clockwise through an angle θ about the unit vector

$$\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix}.$$

Therefore, in 8.4, $a = \cos\phi, b = \sin\phi$, and $c = 0$. It follows the matrix of this transformation with respect to the usual basis is

$$\begin{pmatrix} \cos^2\phi + \sin^2\phi\cos\theta & \cos\phi\sin\phi(1-\cos\theta) & \sin\phi\sin\theta \\ \cos\phi\sin\phi(1-\cos\theta) & \sin^2\phi + \cos^2\phi\cos\theta & -\cos\phi\sin\theta \\ -\sin\phi\sin\theta & \cos\phi\sin\theta & \cos\theta \end{pmatrix} \equiv M_2(\phi,\theta)$$

Finally, we rotate counter clockwise about the positive x_3^2 axis by ψ. The vector in the positive x_3^1 axis is the same as the vector in the fixed x_3 axis. Thus the unit vector in the positive direction of the x_3^2 axis is

$$\begin{pmatrix} \cos^2\phi + \sin^2\phi\cos\theta & \cos\phi\sin\phi(1-\cos\theta) & \sin\phi\sin\theta \\ \cos\phi\sin\phi(1-\cos\theta) & \sin^2\phi + \cos^2\phi\cos\theta & -\cos\phi\sin\theta \\ -\sin\phi\sin\theta & \cos\phi\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \cos^2\phi + \sin^2\phi\cos\theta \\ \cos\phi\sin\phi(1-\cos\theta) \\ -\sin\phi\sin\theta \end{pmatrix} = \begin{pmatrix} \cos^2\phi + \sin^2\phi\cos\theta \\ \cos\phi\sin\phi(1-\cos\theta) \\ -\sin\phi\sin\theta \end{pmatrix}$$

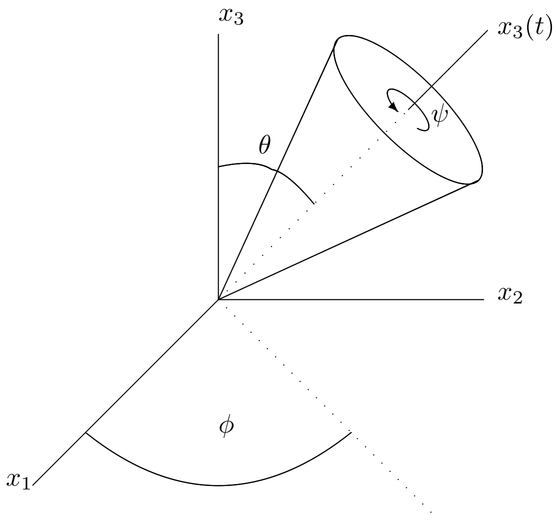
and it is desired to rotate counter clockwise through an angle of ψ about this vector. Thus, in this case,

$$a = \cos^2\phi + \sin^2\phi\cos\theta, b = \cos\phi\sin\phi(1-\cos\theta), c = -\sin\phi\sin\theta.$$

and you could substitute in to the formula of Theorem 8.4 and obtain a matrix which represents the linear transformation obtained by rotating counter clockwise about the positive x_3^2 axis, $M_3(\phi,\theta,\psi)$. Then what would be the matrix with respect to the usual basis for the linear transformation which is obtained as a composition of the three just described? By Theorem 8.3.17, this matrix equals the product of these three,

$$M_3(\phi,\theta,\psi) M_2(\phi,\theta) M_1(\phi).$$

I leave the details to you. There are procedures due to Lagrange which will allow you to write differential equations for the Euler angles in a rotating body. To give an idea how these angles apply, consider the following picture.



This is as far as I will go on this topic. The point is, it is possible to give a systematic description in terms of matrix multiplication of a very elaborate geometrical description of a composition of linear transformations. You see from the picture it is possible to describe the motion of the spinning top shown in terms of these Euler angles.

8.4 Eigenvalues And Eigenvectors Of Linear Transformations

Let V be a finite dimensional vector space. For example, it could be a subspace of $\mathbb{C}^n$or $\mathbb{R}^n$. Also suppose $A \in \mathcal{L}(V,V)$.

Definition 8.4.1 *The characteristic polynomial of* A *is defined as* $q(\lambda) \equiv \det(\lambda I - A)$ *. The zeros of* $q(\lambda)$ *in* $\mathbb{F}$ *are called the eigenvalues of* A*.*

Lemma 8.4.2 *When* λ *is an eigenvalue of* A *which is also in* $\mathbb{F}$*, the field of scalars, then there exists* $v \neq 0$ *such that* $Av = \lambda v$*.*

Proof: This follows from Theorem 8.3.16. Since $\lambda \in \mathbb{F}$,

$$\lambda I - A \in \mathcal{L}(V,V)$$

and since it has zero determinant, it is not one to one. ■

The following lemma gives the existence of something called the minimal polynomial.

Lemma 8.4.3 *Let* $A \in \mathcal{L}(V,V)$ *where* V *is a finite dimensional vector space of dimension* n *with arbitrary field of scalars. Then there exists a unique polynomial of the form*

$$p(\lambda) = \lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0$$

such that $p(A) = 0$ *and* m *is as small as possible for this to occur.*

Proof: Consider the linear transformations, $I, A, A^2, \cdots, A^{n^2}$. There are n^2+1 of these transformations and so by Theorem 8.2.3 the set is linearly dependent. Thus there exist constants, $c_i \in \mathbb{F}$ such that

$$c_0 I + \sum_{k=1}^{n^2} c_k A^k = 0.$$

This implies there exists a polynomial, $q(\lambda)$ which has the property that $q(A) = 0$. In fact, one example is $q(\lambda) \equiv c_0 + \sum_{k=1}^{n^2} c_k\lambda^k$. Dividing by the leading term, it can be assumed this polynomial is of the form $\lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0$, a monic polynomial. Now consider all such monic polynomials, q such that $q(A) = 0$ and pick the one which has the smallest degree m. This is called the minimal polynomial and will be denoted here by $p(\lambda)$. If there were two minimal polynomials, the one just found and another,

$$\lambda^m + d_{m-1}\lambda^{m-1} + \cdots + d_1\lambda + d_0.$$

Then subtracting these would give the following polynomial,

$$\widetilde{q}(\lambda) = (d_{m-1} - c_{m-1})\,\lambda^{m-1} + \cdots + (d_1 - c_1)\,\lambda + d_0 - c_0$$

Since $\widetilde{q}(A) = 0$, this requires each $d_k = c_k$ since otherwise you could divide by $d_k - c_k$ where k is the largest one which is nonzero. Thus the choice of m would be contradicted. ■

Theorem 8.4.4 *Let V be a nonzero finite dimensional vector space of dimension n with the field of scalars equal to $\mathbb{F}$. Suppose $A \in \mathcal{L}(V,V)$ and for $p(\lambda)$ the minimal polynomial defined above, let $\mu \in \mathbb{F}$ be a zero of this polynomial. Then there exists $v \neq 0, v \in V$ such that*

$$Av = \mu v.$$

If $\mathbb{F} = \mathbb{C}$, then A always has an eigenvector and eigenvalue. Furthermore, if $\{\lambda_1, \cdots, \lambda_m\}$ are the zeros of $p(\lambda)$ in $\mathbb{F}$, these are exactly the eigenvalues of A for which there exists an eigenvector in V.

Proof: Suppose first μ is a zero of $p(\lambda)$. Since $p(\mu) = 0$, it follows

$$p(\lambda) = (\lambda - \mu)\,k(\lambda)$$

where $k(\lambda)$ is a polynomial having coefficients in $\mathbb{F}$. Since p has minimal degree, $k(A) \neq 0$ and so there exists a vector, $u \neq 0$ such that $k(A)\,u \equiv v \neq 0$. But then

$$(A - \mu I)\,v = (A - \mu I)\,k(A)\,(u) = \mathbf{0}.$$

The next claim about the existence of an eigenvalue follows from the fundamental theorem of algebra and what was just shown.

It has been shown that every zero of $p(\lambda)$ is an eigenvalue which has an eigenvector in V. Now suppose μ is an eigenvalue which has an eigenvector in V so that $Av = \mu v$ for some $v \in V, v \neq 0$. Does it follow μ is a zero of $p(\lambda)$?

$$\mathbf{0} = p(A)\, v = p(\mu)\, v$$

and so μ is indeed a zero of $p(\lambda)$. ■

In summary, the theorem says that the eigenvalues which have eigenvectors in V are exactly the zeros of the minimal polynomial which are in the field of scalars $\mathbb{F}$.

8.5 Exercises

1. If A, B, and C are each $n \times n$ matrices and ABC is invertible, why are each of A, B, and C invertible?
2. Give an example of a 3×2 matrix with the property that the linear transformation determined by this matrix is one to one but not onto.
3. Explain why $A\mathbf{x} = \mathbf{0}$ always has a solution whenever A is a linear transformation.
4. Review problem: Suppose $\det(A - \lambda I) = 0$. Show using Theorem 3.1.15 there exists $\mathbf{x} \neq \mathbf{0}$ such that $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$.
5. How does the minimal polynomial of an algebraic number relate to the minimal polynomial of a linear transformation? Can an algebraic number be thought of as a linear transformation? How?
6. Recall the fact from algebra that if $p(\lambda)$ and $q(\lambda)$ are polynomials, then there exists $l(\lambda)$, a polynomial such that

$$q(\lambda) = p(\lambda)\, l(\lambda) + r(\lambda)$$

where the degree of $r(\lambda)$ is less than the degree of $p(\lambda)$ or else $r(\lambda) = 0$. With this in mind, why must the minimal polynomial always divide the characteristic polynomial? That is, why does there always exist a polynomial $l(\lambda)$ such that $p(\lambda)\, l(\lambda) = q(\lambda)$? Can you give conditions which imply the minimal polynomial equals the characteristic polynomial? Go ahead and use the Cayley Hamilton theorem.

7. In the following examples, a linear transformation, T is given by specifying its action on a basis β. Find its matrix with respect to this basis.

(a) $T\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1\begin{pmatrix} -1 \\ 1 \end{pmatrix}, T\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

(b) $T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 0 \\ 1 \end{pmatrix} + 1\begin{pmatrix} -1 \\ 1 \end{pmatrix}, T\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

(c) $T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 0 \end{pmatrix}, T\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

8. Let $\beta = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be a basis for $\mathbb{F}^n$ and let $T : \mathbb{F}^n \to \mathbb{F}^n$ be defined as follows.

$$T\left(\sum_{k=1}^{n} a_k \mathbf{u}_k\right) = \sum_{k=1}^{n} a_k b_k \mathbf{u}_k$$

First show that T is a linear transformation. Next show that the matrix of T with respect to this basis, $[T]_\beta$ is

$$\begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix}$$

Show that the above definition is equivalent to simply specifying T on the basis vectors of β by

$$T(\mathbf{u}_k) = b_k \mathbf{u}_k.$$

9. ↑In the situation of the above problem, let $\gamma = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ be the standard basis for $\mathbb{F}^n$ where $\mathbf{e}_k$ is the vector which has 1 in the k^{th} entry and zeros elsewhere. Show that $[T]_\gamma =$

$$\begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} [T]_\beta \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}^{-1} \tag{8.5}$$

10. ↑Generalize the above problem to the situation where T is given by specifying its action on the vectors of a basis $\beta = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ as follows.

$$T\mathbf{u}_k = \sum_{j=1}^{n} a_{jk} \mathbf{u}_j.$$

Letting $A = (a_{ij})$, verify that for $\gamma = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, 8.5 still holds and that $[T]_\beta = A$.

11. Let P_3 denote the set of real polynomials of degree no more than 3, defined on an interval $[a, b]$. Show that P_3 is a subspace of the vector space of all functions defined on this interval. Show that a basis for P_3 is $\{1, x, x^2, x^3\}$. Now let D denote the differentiation operator which sends a function to its derivative. Show D is a linear transformation which sends P_3 to P_3. Find the matrix of this linear transformation with respect to the given basis.

12. Generalize the above problem to P_n, the space of polynomials of degree no more than n with basis $\{1, x, \cdots, x^n\}$.

13. In the situation of the above problem, let the linear transformation be $T = D^2 + 1$, defined as $Tf = f'' + f$. Find the matrix of this linear transformation with respect to the given basis $\{1, x, \cdots, x^n\}$. Write it down for $n = 4$.

14. In calculus, the following situation is encountered. There exists a vector valued function $\mathbf{f} : U \to \mathbb{R}^m$ where U is an open subset of $\mathbb{R}^n$. Such a function is said to have a derivative or to be differentiable at $\mathbf{x} \in U$ if there exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{v} \to \mathbf{0}} \frac{|\mathbf{f}(\mathbf{x} + \mathbf{v}) - \mathbf{f}(\mathbf{x}) - T\mathbf{v}|}{|\mathbf{v}|} = 0.$$

First show that this linear transformation, if it exists, must be unique. Next show that for $\beta = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, , the standard basis, the k^{th} column of $[T]_\beta$ is

$$\frac{\partial \mathbf{f}}{\partial x_k}(\mathbf{x}).$$

Actually, the result of this problem is a well kept secret. People typically don't see this in calculus. It is seen for the first time in advanced calculus if then.

15. Recall that A is similar to B if there exists a matrix P such that $A = P^{-1}BP$. Show that if A and B are similar, then they have the same determinant. Give an example of two matrices which are not similar but have the same determinant.

16. Suppose $A \in \mathcal{L}(V, W)$ where $\dim(V) > \dim(W)$. Show $\ker(A) \neq \{\mathbf{0}\}$. That is, show there exist nonzero vectors $\mathbf{v} \in V$ such that $A\mathbf{v} = \mathbf{0}$.

17. A vector $\mathbf{v}$ is in the convex hull of a nonempty set S if there are finitely many vectors of $S, \{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ and nonnegative scalars $\{t_1, \cdots, t_m\}$ such that

$$\mathbf{v} = \sum_{k=1}^{m} t_k \mathbf{v}_k, \ \sum_{k=1}^{m} t_k = 1.$$

Such a linear combination is called a convex combination. Suppose now that $S \subseteq V$, a vector space of dimension n. Show that if $\mathbf{v} = \sum_{k=1}^{m} t_k \mathbf{v}_k$ is a vector in the convex hull for $m > n + 1$, then there exist other scalars $\{t'_k\}$ such that

$$\mathbf{v} = \sum_{k=1}^{m-1} t'_k \mathbf{v}_k.$$

Thus every vector in the convex hull of S can be obtained as a convex combination of at most $n + 1$ points of S. This incredible result is in Rudin [24]. **Hint:** Consider $L : \mathbb{R}^m \to V \times \mathbb{R}$ defined by

$$L(\mathbf{a}) \equiv \left(\sum_{k=1}^{m} a_k \mathbf{v}_k, \sum_{k=1}^{m} a_k \right)$$

Explain why $\ker(L) \neq \{\mathbf{0}\}$. Next, letting $\mathbf{a} \in \ker(L) \setminus \{\mathbf{0}\}$ and $\lambda \in \mathbb{R}$, note that $\lambda\mathbf{a} \in \ker(L)$. Thus for all $\lambda \in \mathbb{R}$,

$$\mathbf{v} = \sum_{k=1}^{m} (t_k + \lambda a_k) \mathbf{v}_k.$$

Now vary λ till some $t_k + \lambda a_k = 0$ for some $a_k \neq 0$.

18. For those who know about compactness, use Problem 17 to show that if $S \subseteq \mathbb{R}^n$ and S is compact, then so is its convex hull.

19. Suppose $A\mathbf{x} = \mathbf{b}$ has a solution. Explain why the solution is unique precisely when $A\mathbf{x} = \mathbf{0}$ has only the trivial (zero) solution.

20. Let A be an $n \times n$ matrix of elements of $\mathbb{F}$. There are two cases. In the first case, $\mathbb{F}$ contains a splitting field of $p_A(\lambda)$ so that $p(\lambda)$ factors into a product of linear polynomials having coefficients in $\mathbb{F}$. It is the second case which is of interest here where $p_A(\lambda)$ does not factor into linear factors having coefficients in $\mathbb{F}$. Let $\mathbb{G}$ be a splitting field of $p_A(\lambda)$ and let $q_A(\lambda)$ be the minimal polynomial of A with respect to the field $\mathbb{G}$. Explain why $q_A(\lambda)$ must divide $p_A(\lambda)$. Now why must $q_A(\lambda)$ factor completely into linear factors?

21. In Lemma 8.2.2 verify that L is linear.

Chapter 9

Canonical Forms

9.1 A Theorem Of Sylvester, Direct Sums

The notation is defined as follows.

Definition 9.1.1 *Let* $L \in \mathcal{L}(V, W)$. *Then* $\ker(L) \equiv \{v \in V : Lv = 0\}$.

Lemma 9.1.2 *Whenever* $L \in \mathcal{L}(V, W)$, $\ker(L)$ *is a subspace.*

Proof: If a, b are scalars and v,w are in $\ker(L)$, then

$$L(av + bw) = aL(v) + bL(w) = 0 + 0 = 0 \blacksquare$$

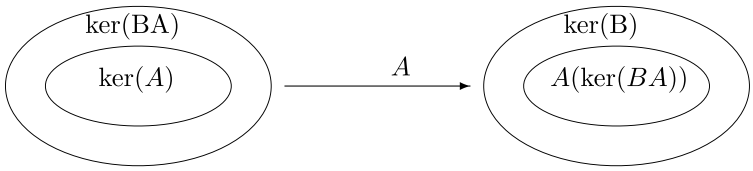
Suppose now that $A \in \mathcal{L}(V, W)$ and $B \in \mathcal{L}(W, U)$ where V, W, U are all finite dimensional vector spaces. Then it is interesting to consider $\ker(BA)$. The following theorem of Sylvester is a very useful and important result.

Theorem 9.1.3 *Let* $A \in \mathcal{L}(V, W)$ *and* $B \in \mathcal{L}(W, U)$ *where* V, W, U *are all vector spaces over a field* $\mathbb{F}$. *Suppose also that* $\ker(A)$ *and* $A(\ker(BA))$ *are finite dimensional subspaces. Then*

$$\dim(\ker(BA)) \leq \dim(\ker(B)) + \dim(\ker(A)).$$

Equality holds if and only if $A(\ker(BA)) = \ker(B)$.

Proof: If $\mathbf{x} \in \ker(BA)$, then $A\mathbf{x} \in \ker(B)$ and so $A(\ker(BA)) \subseteq \ker(B)$. The following picture may help.



Now let $\{x_1, \cdots, x_n\}$ be a basis of $\ker(A)$ and let $\{Ay_1, \cdots, Ay_m\}$ be a basis for $A(\ker(BA))$. Take any $z \in \ker(BA)$. Then $Az = \sum_{i=1}^{m} a_i Ay_i$ and so

$$A\left(z - \sum_{i=1}^{m} a_i y_i\right) = \mathbf{0}$$

which means $z - \sum_{i=1}^{m} a_i y_i \in \ker(A)$ and so there are scalars b_i such that

$$z - \sum_{i=1}^{m} a_i y_i = \sum_{j=1}^{n} b_i x_i.$$

It follows $\text{span}(x_1, \cdots, x_n, y_1, \cdots, y_m) \supseteq \ker(BA)$ and so by the first part, (See the picture.)

$$\dim(\ker(BA)) \leq n + m \leq \dim(\ker(A)) + \dim(\ker(B))$$

Now $\{x_1, \cdots, x_n, y_1, \cdots, y_m\}$ is linearly independent because if

$$\sum_i a_i x_i + \sum_j b_j y_j = 0$$

then you could do A to both sides and conclude that $\sum_j b_j A y_j = 0$ which requires that each $b_j = 0$. Then it follows that each $a_i = 0$ also because it implies $\sum_i a_i x_i = 0$. Thus

$$\{x_1, \cdots, x_n, y_1, \cdots, y_m\}$$

is a basis for $\ker(BA)$. Then $A(\ker(BA)) = \ker(B)$ if and only if $m = \dim(\ker(B))$ if and only if

$$\dim(\ker(BA)) = m + n = \dim(\ker(B)) + \dim(\ker(A)). \blacksquare$$

Of course this result holds for any finite product of linear transformations by induction. One way this is quite useful is in the case where you have a finite product of linear transformations $\prod_{i=1}^{l} L_i$ all in $\mathcal{L}(V, V)$. Then

$$\dim\left(\ker \prod_{i=1}^{l} L_i\right) \leq \sum_{i=1}^{l} \dim(\ker L_i).$$

Definition 9.1.4 *Let $\{V_i\}_{i=1}^r$ be subspaces of V. Then*

$$\sum_{i=1}^{r} V_i \equiv V_1 + \cdots + V_r$$

denotes all sums of the form $\sum_{i=1}^r v_i$ where $v_i \in V_i$. If whenever

$$\sum_{i=1}^{r} v_i = 0, v_i \in V_i, \tag{9.1}$$

it follows that $v_i = 0$ for each i, then a special notation is used to denote $\sum_{i=1}^r V_i$. This notation is

$$V_1 \oplus \cdots \oplus V_r,$$

and it is called a direct sum of subspaces.

Now here is a useful lemma which is likely already understood.

Lemma 9.1.5 *Let $L \in \mathcal{L}(V, W)$ where V, W are n dimensional vector spaces. Then L is one to one, if and only if L is also onto. In fact, if $\{v_1, \cdots, v_n\}$ is a basis, then so is $\{Lv_1, \cdots, Lv_n\}$.*

Proof: Let $\{v_1, \cdots, v_n\}$ be a basis for V. Then I claim that $\{Lv_1, \cdots, Lv_n\}$ is a basis for W. First of all, I show $\{Lv_1, \cdots, Lv_n\}$ is linearly independent. Suppose

$$\sum_{k=1}^{n} c_k L v_k = 0.$$

Then

$$L\left(\sum_{k=1}^{n} c_k v_k\right) = 0$$

and since L is one to one, it follows

$$\sum_{k=1}^{n} c_k v_k = 0$$

which implies each $c_k = 0$. Therefore, $\{Lv_1, \cdots, Lv_n\}$ is linearly independent. If there exists w not in the span of these vectors, then by Lemma 7.2.10, $\{Lv_1, \cdots, Lv_n, w\}$ would be independent and this contradicts the exchange theorem, Theorem 7.2.4 because it would be a linearly independent set having more vectors than the spanning set $\{v_1, \cdots, v_n\}$.

Conversely, suppose L is onto. Then there exists a basis for W which is of the form $\{Lv_1, \cdots, Lv_n\}$. It follows that $\{v_1, \cdots, v_n\}$ is linearly independent. Hence it is a basis for V by similar reasoning to the above. Then if $Lx = 0$, it follows that there are scalars c_i such that $x = \sum_i c_i v_i$ and consequently $0 = Lx = \sum_i c_i L v_i$. Therefore, each $c_i = 0$ and so $x = 0$ also. Thus L is one to one. ■

Lemma 9.1.6 *If $V = V_1 \oplus \cdots \oplus V_r$ and if $\beta_i = \left\{v_1^i, \cdots, v_{m_i}^i\right\}$ is a basis for V_i, then a basis for V is $\{\beta_1, \cdots, \beta_r\}$. Thus*

$$\dim(V) = \sum_{i=1}^{r} \dim(V_i).$$

Proof: Suppose $\sum_{i=1}^{r} \sum_{j=1}^{m_i} c_{ij} v_j^i = 0$. then since it is a direct sum, it follows for each i,

$$\sum_{j=1}^{m_i} c_{ij} v_j^i = 0$$

and now since $\left\{v_1^i, \cdots, v_{m_i}^i\right\}$ is a basis, each $c_{ij} = 0$. ■

Here is a fundamental lemma.

Lemma 9.1.7 *Let L_i be in $\mathcal{L}(V, V)$ and suppose for $i \neq j, L_i L_j = L_j L_i$ and also L_i is one to one on $\ker(L_j)$ whenever $i \neq j$. Then*

$$\ker\left(\prod_{i=1}^{p} L_i\right) = \ker(L_1) \oplus + \cdots + \oplus \ker(L_p)$$

Here $\prod_{i=1}^{p} L_i$ is the product of all the linear transformations.

Proof : Note that since the operators commute, $L_j : \ker(L_i) \to \ker(L_i)$. Here is why. If $L_i y = 0$ so that $y \in \ker(L_i)$, then

$$L_i L_j y = L_j L_i y = L_j 0 = 0$$

and so $L_j : \ker(L_i) \mapsto \ker(L_i)$. Next observe that it is obvious that, since the operators commute,

$$\sum_{i=1}^{p} \ker(L_p) \subseteq \ker\left(\prod_{i=1}^{p} L_i\right)$$

Next, why is $\sum_i \ker(L_p) = \ker(L_1) \oplus \cdots \oplus \ker(L_p)$? Suppose

$$\sum_{i=1}^{p} v_i = 0,\ v_i \in \ker(L_i),$$

but some $v_i \neq 0$. Then do $\prod_{j\neq i} L_j$ to both sides. Since the linear transformations commute, this results in

$$\prod_{j\neq i} L_j(v_i) = 0$$

which contradicts the assumption that these L_j are one to one on $\ker(L_i)$ and the observation that they map $\ker(L_i)$ to $\ker(L_i)$. Thus if

$$\sum_{i} v_i = 0,\ v_i \in \ker(L_i)$$

then each $v_i = 0$. It follows that

$$\ker(L_1) \oplus + \cdots + \oplus \ker(L_p) \subseteq \ker\left(\prod_{i=1}^{p} L_i\right) \qquad (*)$$

From Sylvester's theorem and the observation about direct sums in Lemma 9.1.6,

$$\begin{aligned}\sum_{i=1}^{p} \dim(\ker(L_i)) &= \dim(\ker(L_1) \oplus + \cdots + \oplus \ker(L_p)) \\ &\leq \dim\left(\ker\left(\prod_{i=1}^{p} L_i\right)\right) \leq \sum_{i=1}^{p} \dim(\ker(L_i))\end{aligned}$$

which implies all these are equal. Now in general, if W is a subspace of V, a finite dimensional vector space and the two have the same dimension, then $W = V$. This is because W has a basis and if v is not in the span of this basis, then v adjoined to the basis of W would be a linearly independent set so the dimension of V would then be strictly larger than the dimension of W. It follows from * that

$$\ker(L_1) \oplus + \cdots + \oplus \ker(L_p) = \ker\left(\prod_{i=1}^{p} L_i\right) \ \blacksquare$$

9.2 Direct Sums, Block Diagonal Matrices

Let V be a finite dimensional vector space with field of scalars $\mathbb{F}$. Here I will make no assumption on $\mathbb{F}$. Also suppose $A \in \mathcal{L}(V,V)$.

Recall Lemma 8.4.3 which gives the existence of the minimal polynomial for a linear transformation A. This is the monic polynomial p which has smallest possible degree such that $p(A) = 0$. It is stated again for convenience.

Lemma 9.2.1 *Let $A \in \mathcal{L}(V,V)$ where V is a finite dimensional vector space of dimension n with field of scalars $\mathbb{F}$. Then there exists a unique monic polynomial of the form*

$$p(\lambda) = \lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0$$

such that $p(A) = 0$ and m is as small as possible for this to occur.

Now it is time to consider the notion of a direct sum of subspaces. Recall you can always assert the existence of a factorization of the minimal polynomial into a product of irreducible polynomials. This fact will now be used to show how to obtain such a direct sum of subspaces.

Definition 9.2.2 *For* $A \in \mathcal{L}(V,V)$ *where* $\dim(V) = n$*, suppose the minimal polynomial is*

$$p(\lambda) = \prod_{k=1}^{q} (\phi_k(\lambda))^{r_k}$$

where the polynomials ϕ_k *have coefficients in* $\mathbb{F}$ *and are irreducible. Now define the generalized eigenspaces*

$$V_k \equiv \ker((\phi_k(A))^{r_k})$$

Note that if one of these polynomials $(\phi_k(\lambda))^{r_k}$ *is a monic linear polynomial, then the generalized eigenspace would be an eigenspace.*

Theorem 9.2.3 *In the context of Definition 9.2.2,*

$$V = V_1 \oplus \cdots \oplus V_q \tag{9.2}$$

and each V_k *is* A *invariant, meaning* $A(V_k) \subseteq V_k$*.* $\phi_l(A)$ *is one to one on each* V_k *for* $k \neq l$*. If* $\beta_i = \{v_1^i, \cdots, v_{m_i}^i\}$ *is a basis for* V_i*, then* $\{\beta_1, \beta_2, \cdots, \beta_q\}$ *is a basis for* V*.*

Proof: It is clear V_k is a subspace which is A invariant because A commutes with $\phi_k(A)^{m_k}$. It is clear the operators $\phi_k(A)^{r_k}$ commute. Thus if $v \in V_k$,

$$\phi_k(A)^{r_k} \phi_l(A)^{r_l} v = \phi_l(A)^{r_l} \phi_k(A)^{r_k} v = \phi_l(A)^{r_l} 0 = 0$$

and so $\phi_l(A)^{r_l} : V_k \to V_k$.

I claim $\phi_l(A)$ is one to one on V_k whenever $k \neq l$. The two polynomials $\phi_l(\lambda)$ and $\phi_k(\lambda)^{r_k}$ are relatively prime so there exist polynomials $m(\lambda), n(\lambda)$ such that

$$m(\lambda)\phi_l(\lambda) + n(\lambda)\phi_k(\lambda)^{r_k} = 1$$

It follows that the sum of all coefficients of λ raised to a positive power are zero and the constant term on the left is 1. Therefore, using the convention $A^0 = I$ it follows

$$m(A)\phi_l(A) + n(A)\phi_k(A)^{r_k} = I$$

If $v \in V_k$, then from the above,

$$m(A)\phi_l(A)v + n(A)\phi_k(A)^{r_k} v = v$$

Since v is in V_k, it follows by definition,

$$m(A)\phi_l(A)v = v$$

and so $\phi_l(A)v \neq 0$ unless $v = 0$. Thus $\phi_l(A)$ and hence $\phi_l(A)^{r_l}$ is one to one on V_k for every $k \neq l$. By Lemma 9.1.7 and the fact that $\ker\left(\prod_{k=1}^{q} \phi_k(\lambda)^{r_k}\right) = V$, 9.2 is obtained. The claim about the bases follows from Lemma 9.1.6. ■

You could consider the restriction of A to V_k. It turns out that this restriction has minimal polynomial equal to $\phi_k(\lambda)^{m_k}$.

Corollary 9.2.4 *Let the minimal polynomial of* A *be* $p(\lambda) = \prod_{k=1}^{q} \phi_k(\lambda)^{m_k}$ *where each* ϕ_k *is irreducible. Let* $V_k = \ker(\phi(A)^{m_k})$*. Then*

$$V_1 \oplus \cdots \oplus V_q = V$$

and letting A_k *denote the restriction of* A *to* V_k*, it follows the minimal polynomial of* A_k *is* $\phi_k(\lambda)^{m_k}$*.*

Proof: Recall the direct sum, $V_1 \oplus \cdots \oplus V_q = V$ where $V_k = \ker(\phi_k(A)^{m_k})$ for $p(\lambda) = \prod_{k=1}^{q} \phi_k(\lambda)^{m_k}$ the minimal polynomial for A where the $\phi_k(\lambda)$ are all irreducible. Thus each V_k is invariant with respect to A. What is the minimal polynomial of A_k, the restriction of A to V_k? First note that $\phi_k(A_k)^{m_k}(V_k) = \{0\}$ by definition. Thus if $\eta(\lambda)$ is the minimal polynomial for A_k then it must divide $\phi_k(\lambda)^{m_k}$ and so by Corollary 7.3.11 $\eta(\lambda) = \phi_k(\lambda)^{r_k}$ where $r_k \leq m_k$. Could $r_k < m_k$? No, this is not possible because then $p(\lambda)$ would fail to be the minimal polynomial for A. You could substitute for the term $\phi_k(\lambda)^{m_k}$ in the factorization of $p(\lambda)$ with $\phi_k(\lambda)^{r_k}$ and the resulting polynomial p' would satisfy $p'(A) = 0$. Here is why. From Theorem 9.2.3, a typical $x \in V$ is of the form

$$\sum_{i=1}^{q} v_i,\ v_i \in V_i$$

Then since all the factors commute,

$$p'(A)\left(\sum_{i=1}^{q} v_i\right) = \prod_{i\neq k}^{q} \phi_i(A)^{m_i} \phi_k(A)^{r_k} \left(\sum_{i=1}^{q} v_i\right)$$

For $j \neq k$

$$\prod_{i\neq k}^{q} \phi_i(A)^{m_i} \phi_k(A)^{r_k} v_j = \prod_{i\neq k,j}^{q} \phi_i(A)^{m_i} \phi_k(A)^{r_k} \phi_j(A)^{m_j} v_j = 0$$

If $j = k$,

$$\prod_{i\neq k}^{q} \phi_i(A)^{m_i} \phi_k(A)^{r_k} v_k = 0$$

which shows $p'(\lambda)$ is a monic polynomial having smaller degree than $p(\lambda)$ such that $p'(A) = 0$. Thus the minimal polynomial for A_k is $\phi_k(\lambda)^{m_k}$ as claimed. ■

How does Theorem 9.2.3 relate to matrices?

Theorem 9.2.5 *Suppose V is a vector space with field of scalars $\mathbb{F}$ and $A \in \mathcal{L}(V,V)$. Suppose also*

$$V = V_1 \oplus \cdots \oplus V_q$$

where each V_k is A invariant. ($AV_k \subseteq V_k$) Also let β_k be an ordered basis for V_k and let A_k denote the restriction of A to V_k. Letting M^k denote the matrix of A_k with respect to this basis, it follows the matrix of A with respect to the basis $\{\beta_1, \cdots, \beta_q\}$ is

$$\begin{pmatrix} M^1 & & 0 \\ & \ddots & \\ 0 & & M^q \end{pmatrix}$$

Proof: Let β denote the ordered basis $\{\beta_1, \cdots, \beta_q\}$, $|\beta_k|$ being the number of vectors in β_k. Let $q_k : \mathbb{F}^{|\beta_k|} \to V_k$ be the usual map such that the following diagram commutes.

$$\begin{array}{ccc} & A_k & \\ V_k & \to & V_k \\ q_k \uparrow & \circ & \uparrow q_k \\ \mathbb{F}^{|\beta_k|} & \to & \mathbb{F}^{|\beta_k|} \\ & M^k & \end{array}$$

Thus $A_k q_k = q_k M^k$. Then if q is the map from $\mathbb{F}^n$ to V corresponding to the ordered basis β just described,

$$q\begin{pmatrix} \mathbf{0} & \cdots & \mathbf{x} & \cdots & \mathbf{0} \end{pmatrix}^T = q_k \mathbf{x},$$

where $\mathbf{x}$ occupies the positions between $\sum_{i=1}^{k-1} |\beta_i| + 1$ and $\sum_{i=1}^{k} |\beta_i|$. Then M will be the matrix of A with respect to β if and only if a similar diagram to the above commutes. Thus it is required that $Aq = qM$. However, from the description of q just made, and the invariance of each V_k,

$$Aq\begin{pmatrix}\mathbf{0}\\ \vdots\\ \mathbf{x}\\ \vdots\\ \mathbf{0}\end{pmatrix}=A_kq_k\mathbf{x}=q_kM^k\mathbf{x}=q\begin{pmatrix}M^1 & & & & 0\\ & \ddots & & & \\ & & M^k & & \\ & & & \ddots & \\ 0 & & & & M^q\end{pmatrix}\begin{pmatrix}\mathbf{0}\\ \vdots\\ \mathbf{x}\\ \vdots\\ \mathbf{0}\end{pmatrix}$$

It follows that the above block diagonal matrix is the matrix of A with respect to the given ordered basis. ■

An examination of the proof of the above theorem yields the following corollary.

Corollary 9.2.6 *If any β_k in the above consists of eigenvectors, then M^k is a diagonal matrix having the corresponding eigenvalues down the diagonal.*

It follows that it would be interesting to consider special bases for the vector spaces in the direct sum. This leads to the Jordan form or more generally other canonical forms such as the rational canonical form.

9.3 Cyclic Sets

It was shown above that for $A \in \mathcal{L}(V,V)$ for V a finite dimensional vector space over the field of scalars $\mathbb{F}$, there exists a direct sum decomposition

$$V = V_1 \oplus \cdots \oplus V_q$$

where

$$V_k = \ker\left(\phi_k(A)^{m_k}\right)$$

and $\phi_k(\lambda)$ is an irreducible polynomial. Here the minimal polynomial of A was

$$\prod_{k=1}^{q}\phi_k(\lambda)^{m_k}$$

Next I will consider the problem of finding a basis for V_k such that the matrix of A restricted to V_k assumes various forms.

Definition 9.3.1 *Letting $x \neq 0$ denote by β_x the vectors $\left\{x, Ax, A^2x, \cdots, A^{m-1}x\right\}$ where m is the smallest such that $A^m x \in \operatorname{span}\left(x, \cdots, A^{m-1}x\right)$. This is called an A cyclic set. The vectors which result are also called a Krylov sequence. For such a sequence of vectors, $|\beta_x| \equiv m$.*

The first thing to notice is that such a Krylov sequence is always linearly independent.

Lemma 9.3.2 *Let $\beta_x = \left\{x, Ax, A^2x, \cdots, A^{m-1}x\right\}, x \neq 0$ where m is the smallest such that $A^m x \in \operatorname{span}\left(x, \cdots, A^{m-1}x\right)$. Then $\left\{x, Ax, A^2x, \cdots, A^{m-1}x\right\}$ is linearly independent.*

Proof: Suppose that there are scalars a_k, not all zero such that

$$\sum_{k=0}^{m-1} a_k A^k x = 0$$

Then letting a_r be the last nonzero scalar in the sum, you can divide by a_r and solve for $A^r x$ as a linear combination of the $A^j x$ for $j < r \leq m-1$ contrary to the definition of m. ■

Now here is a nice lemma which has been pretty much discussed earlier.

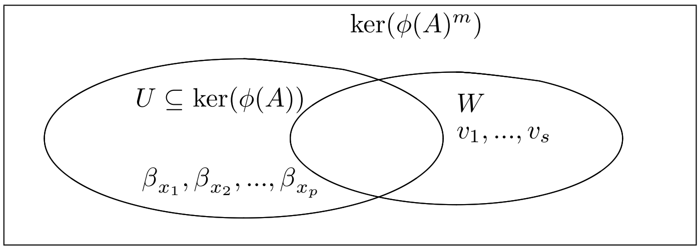
Lemma 9.3.3 *Suppose W is a subspace of V where V is a finite dimensional vector space and $L \in \mathcal{L}(V,V)$ and suppose $LW = LV$. Then $V = W + \ker(L)$.*

Proof: Let a basis for $LV = LW$ be $\{Lw_1, \cdots, Lw_m\}, w_i \in W$. Then let $y \in V$. Thus $Ly = \sum_{i=1}^{m} c_i L w_i$ and so

$$L\left(\overbrace{y - \sum_{i=1}^{m} c_i w_i}^{=z}\right) \equiv Lz = 0$$

It follows that $z \in \ker(L)$ and so $y = \sum_{i=1}^{m} c_i w_i + z \in W + \ker(L)$. ■

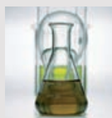
For more on the next lemma and the following theorem, see Hofman and Kunze [15]. I am following the presentation in Friedberg Insel and Spence [10]. See also Herstein [14] for a different approach to canonical forms. To help organize the ideas in the lemma, here is a diagram.



Lemma 9.3.4 *Let W be an A invariant $(AW \subseteq W)$ subspace of $\ker(\phi(A)^m)$ for m a positive integer where $\phi(\lambda)$ is an irreducible monic polynomial of degree d. Let U be an A invariant subspace of $\ker(\phi(A))$.*

If $\{v_1, \cdots, v_s\}$ is a basis for W then if $x \in U \setminus W$,

$$\{v_1, \cdots, v_s, \beta_x\}$$



is linearly independent.

There exist vectors $x_1, \cdots, x_p$ *each in* U *such that*

$$\left\{v_1, \cdots, v_s, \beta_{x_1}, \cdots, \beta_{x_p}\right\}$$

is a basis for

$$U + W.$$

Also, if $x \in \ker(\phi(A)^m)$, $|\beta_x| = kd$ *where* $k \leq m$. *Here* $|\beta_x|$ *is the length of* β_x, *the degree of the monic polynomial* $\eta(\lambda)$ *satisfying* $\eta(A)x = 0$ *with* $\eta(\lambda)$ *having smallest possible degree.*

Proof: Claim: If $x \in \ker \phi(A)$, and $|\beta_x|$ denotes the length of β_x, then $|\beta_x| = d$ the degree of the irreducible polynomial $\phi(\lambda)$ and so

$$\beta_x = \left\{x, Ax, A^2x, \cdots, A^{d-1}x\right\}$$

also span (β_x) is A invariant, $A(\text{span}(\beta_x)) \subseteq \text{span}(\beta_x)$.

Proof of the claim: Let $m = |\beta_x|$. That is, there exists monic $\eta(\lambda)$ of degree m and $\eta(A)x = 0$ with m is as small as possible for this to happen. Then from the usual process of division of polynomials, there exist $l(\lambda), r(\lambda)$ such that $r(\lambda) = 0$ or else has smaller degree than that of $\eta(\lambda)$ such that

$$\phi(\lambda) = \eta(\lambda) l(\lambda) + r(\lambda)$$

If $\deg(r(\lambda)) < \deg(\eta(\lambda))$, then the equation implies $0 = \phi(A)x = r(A)x$ and so m was incorrectly chosen. Hence $r(\lambda) = 0$ and so if $l(\lambda) \neq 1$, then $\eta(\lambda)$ divides $\phi(\lambda)$ contrary to the assumption that $\phi(\lambda)$ is irreducible. Hence $l(\lambda) = 1$ and $\eta(\lambda) = \phi(\lambda)$. The claim about span (β_x) is obvious because $A^dx \in \text{span}(\beta_x)$. This shows the claim.

Suppose now $x \in U \setminus W$ where $U \subseteq \ker(\phi(A))$. Consider

$$\{v_1, \cdots, v_s, \beta_x\}.$$

Is this set of vectors independent? Suppose

$$\sum_{i=1}^{s} a_i v_i + \sum_{j=1}^{d} d_j A^{j-1}x = 0.$$

If $z \equiv \sum_{j=1}^{d} d_j A^{j-1}x$, then $z \in W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right)$. Then also for each $m \leq d-1$,

$$A^m z \in W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right)$$

because W, span $\left(x, Ax, \cdots, A^{d-1}x\right)$ are A invariant. Therefore,

$$\begin{aligned} \text{span}\left(z, Az, \cdots, A^{d-1}z\right) &\subseteq W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right) \\ &\subseteq \text{span}\left(x, Ax, \cdots, A^{d-1}x\right) \end{aligned} \tag{9.3}$$

Suppose $z \neq 0$. Then from the Lemma 9.3.2 above, $\left\{z, Az, \cdots, A^{d-1}z\right\}$ must be linearly independent. Therefore,

$$d = \dim\left(\text{span}\left(z, Az, \cdots, A^{d-1}z\right)\right) \leq \dim\left(W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right)\right)$$

$$\leq \dim\left(\text{span}\left(x, Ax, \cdots, A^{d-1}x\right)\right) = d$$

Thus

$$W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right) = \text{span}\left(x, Ax, \cdots, A^{d-1}x\right)$$

which would require $x \in W$ but this is assumed not to take place. Hence $z = 0$ and so the linear independence of the $\{v_1, \cdots, v_s\}$ implies each $a_i = 0$. Then the linear independence of $\left\{x, Ax, \cdots, A^{d-1}x\right\}$, which follows from Lemma 9.3.2, shows each $d_j = 0$. Thus $\left\{v_1, \cdots, v_s, x, Ax, \cdots, A^{d-1}x\right\}$ is linearly independent as claimed.

Let $x \in U \setminus W \subseteq \ker(\phi(A))$. Then it was just shown that $\{v_1, \cdots, v_s, \beta_x\}$ is linearly independent. Let W_1 be given by

$$y \in \operatorname{span}(v_1, \cdots, v_s, \beta_x) \equiv W_1$$

Then W_1 is A invariant. If W_1 equals $U + W$, then you are done. If not, let W_1 play the role of W and pick $x_1 \in U \setminus W_1$ and repeat the argument. Continue till

$$\operatorname{span}(v_1, \cdots, v_s, \beta_{x_1}, \cdots, \beta_{x_n}) = U + W$$

The process stops because $\ker(\phi(A)^m)$ is finite dimensional.

Finally, letting $x \in \ker(\phi(A)^m)$, there is a monic polynomial $\eta(\lambda)$ such that $\eta(A)x = 0$ and $\eta(\lambda)$ is of smallest possible degree, which degree equals $|\beta_x|$. Then

$$\phi(\lambda)^m = \eta(\lambda) l(\lambda) + r(\lambda)$$

If $\deg(r(\lambda)) < \deg(\eta(\lambda))$, then $r(A)x = 0$ and $\eta(\lambda)$ was incorrectly chosen. Hence $r(\lambda) = 0$ and so $\eta(\lambda)$ must divide $\phi(\lambda)^m$. Hence by Corollary 7.3.11 $\eta(\lambda) = \phi(\lambda)^k$ where $k \leq m$. Thus $|\beta_x| = kd = \deg(\eta(\lambda))$. ■

With this preparation, here is the main result about a basis V where $A \in \mathcal{L}(V,V)$ and the minimal polynomial for A is $\phi(A)^m$ for $\phi(\lambda)$ irreducible an irreducible monic polynomial. There is a very interesting generalization of this theorem in [15] which pertains to the existence of complementary subspaces. For an outline of this generalization, see Problem 9 on Page 293.

Theorem 9.3.5 *Suppose $A \in \mathcal{L}(V,V)$ for V some finite dimensional vector space. Then for each $k \in \mathbb{N}$, there exists a cyclic basis for $\ker\left(\phi(A)^k\right)$ which is one of the form $\beta = \left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$ or $\ker\left(\phi(A)^k\right) = \{0\}$. Note that if $\ker(\phi(A)) \neq \{0\}$, then the same is true for all $\ker\left(\phi(A)^k\right), k \in \mathbb{N}$.*

Proof: If $k = 1$, you can use Lemma 9.3.4 and let $W = \{0\}$ and $U = \ker(\phi(A))$ to obtain the cyclic basis. Suppose then that the theorem is true for $m-1, m-1 \geq 1$ meaning that for any finite dimensional vector space V and $A \in \mathcal{L}(V,V)$, $\ker\left(\phi(A)^k\right)$ has a cyclic basis for all $k \leq m-1$. Consider a new vector space $\phi(A)\ker(\phi(A)^m) \equiv \hat{V}$ in place of V and the restriction of A to $\hat{V}$ which we will call $\hat{A}$. Then $\hat{A} \in \mathcal{L}\left(\hat{V}, \hat{V}\right)$. It follows $\phi(A)^{m-1}(\phi(A)\ker(\phi(A)^m)) = \phi(A)^{m-1}\hat{V} = 0$ and since $\phi(\lambda)$ is irreducible, the minimum polynomial of $\hat{A}$ on $\hat{V}$ is $\phi\left(\hat{A}\right)^k$ for some $k \leq m-1$. Thus $\ker\left(\phi\left(\hat{A}\right)^k\right) \equiv \left\{v \in \hat{V} : \phi\left(\hat{A}\right)^k v = 0\right\}$. Since $k \leq m-1$ the cyclic basis in $\hat{V}$ exists by induction. If $k = 0$, then you would have $\hat{V} = \{0\}$ and $\{0\} = \phi(A)\ker(\phi(A)^m) \supseteq \ker(\phi(A))$ so nothing is of any interest because all of these spaces are $\{0\}$.

Let the cyclic basis for $\hat{V} \equiv \phi(A)\ker(\phi(A)^m)$ be $\left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}, x_i \in \phi(A)\ker(\phi(A)^m)$. Let $x_i = \phi(A)y_i, y_i \in \ker(\phi(A)^m)$. Consider $\left\{\beta_{y_1}, \cdots, \beta_{y_p}\right\}, y_i \in \ker(\phi(A)^m)$. Are these vectors independent? Suppose

$$0 = \sum_{i=1}^{p} \sum_{j=1}^{|\beta_{y_i}|} a_{ij} A^{j-1} y_i \equiv \sum_{i=1}^{p} f_i(A) y_i \tag{9.4}$$

If the sum involved x_i in place of y_i, then something could be said because $\left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$ is a basis.

Do $\phi(A)$ to both sides to obtain

$$0 = \sum_{i=1}^{p} \sum_{j=1}^{|\beta_{y_i}|} a_{ij} A^{j-1} x_i \equiv \sum_{i=1}^{p} f_i\left(\hat{A}\right) x_i$$

Now $f_i\left(\hat{A}\right)x_i = 0$ for each i since $f_i\left(\hat{A}\right)x_i \in \text{span}\left(\beta_{x_i}\right)$ and as just mentioned, $\left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$ is a basis. Let $\eta_i(\lambda)$ be the monic polynomial of smallest degree such that $\eta_i\left(\hat{A}\right)x_i = 0$. Then

$$f_i(\lambda) = \eta_i(\lambda)\, l(\lambda) + r(\lambda)$$

where $r(\lambda) = 0$ or else it has smaller degree than $\eta_i(\lambda)$. However, the equation then shows that $r\left(\hat{A}\right)x_i = 0$ which would contradict the choice of $\eta_i(\lambda)$. Thus $r(\lambda) = 0$ and $\eta_i(\lambda)$ divides $f_i(\lambda)$. Also, $\phi\left(\hat{A}\right)^{m-1}x_i = \phi\left(\hat{A}\right)^{m-1}\phi(A)\,y_i = 0$ and so $\eta_i(\lambda)$ must divide $\phi(\lambda)^{m-1}$. From Corollary 7.3.11, it follows that, since $\phi(\lambda)$ is irreducible, $\eta_i(\lambda) = \phi(\lambda)^r$ for some $r \leq m-1$. Thus $\phi(\lambda)$ divides $\eta_i(\lambda)$ which divides $f_i(\lambda)$. Hence $f_i(\lambda) = \phi(\lambda)\, g_i(\lambda)$! Now

$$0 = \sum_{i=1}^{p} f_i(A)\, y_i = \sum_{i=1}^{p} g_i(A)\,\phi(A)\, y_i = \sum_{i=1}^{p} g_i\left(\hat{A}\right)x_i.$$

By the same reasoning just given, since $g_i\left(\hat{A}\right)x_i \in \text{span}\left(\beta_{x_i}\right)$, it follows that each $g_i\left(\hat{A}\right)x_i = 0$. Therefore,

$$f_i(A)\, y_i = g_i\left(\hat{A}\right)\phi(A)\, y_i = g_i\left(\hat{A}\right)x_i = 0.$$

Therefore,

$$f_i(A)\, y_i = \sum_{j=1}^{\left|\beta_{y_j}\right|} a_{ij}A^{j-1}y_i = 0$$

and by independence of the β_{y_i}, this implies $a_{ij} = 0$ for each j for each i.

Next, it follows from the definition that $\phi(A)\left(\ker\left(\phi(A)^m\right)\right) = \text{span}\left(\beta_{x_1}, \cdots, \beta_{x_p}\right)$.

Now

$$W \equiv \text{span}\left(\beta_{y_1}, \cdots, \beta_{y_p}\right) \subseteq \ker\left(\phi\left(A\right)^m\right)$$

because each $y_i \in \ker\left(\phi\left(A\right)^m\right)$. Then from the above description of $\left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$ as a cyclic basis for $\phi\left(A\right)\left(\ker\left(\phi\left(A\right)^m\right)\right)$,

$$\begin{aligned}\phi\left(A\right)\left(\ker\left(\phi\left(A\right)^m\right)\right) &= \text{span}\left(\beta_{x_1}, \cdots, \beta_{x_p}\right) \subseteq \phi\left(A\right)\text{span}\left(\beta_{y_1}, \cdots, \beta_{y_p}\right) \\ &\equiv \phi\left(A\right)\left(W\right) \subseteq \phi\left(A\right)\ker\left(\phi\left(A\right)^m\right)\end{aligned}$$

To see the first inclusion,

$$A^r x_q = A^r \phi\left(A\right) y_q = \phi\left(A\right) A^r y_q \in \phi\left(A\right)\text{span}\left(\beta_{y_q}\right) \subseteq \phi\left(A\right)\text{span}\left(\beta_{y_1}, \cdots, \beta_{y_p}\right)$$

It follows from Lemma 9.3.3 that $\ker\left(\phi\left(A\right)^m\right) = W + \ker\left(\phi\left(A\right)\right)$. From Lemma 9.3.4 $W + \ker\left(\phi\left(A\right)\right)$ has a basis of the form $\left\{\beta_{y_1}, \cdots, \beta_{y_p}, \beta_{z_1}, \cdots, \beta_{z_s}\right\}$. ■

9.4 Nilpotent Transformations

Definition 9.4.1 *Let V be a vector space over the field of scalars $\mathbb{F}$. Then $N \in \mathcal{L}\left(V, V\right)$ is called nilpotent if for some m, it follows that $N^m = 0$.*

The following lemma contains some significant observations about nilpotent transformations.

Lemma 9.4.2 *Suppose $N^k x \neq 0$. Then $\left\{x, Nx, \cdots, N^k x\right\}$ is linearly independent. Also, the minimal polynomial of N is λ^m where m is the first such that $N^m = 0$.*

Proof: Suppose $\sum_{i=0}^{k} c_i N^i x = 0$ where not all $c_i = 0$. There exists l such that $k \leq l < m$ and $N^{l+1} x = 0$ but $N^l x \neq 0$. Then multiply both sides by N^l to conclude that $c_0 = 0$. Next multiply both sides by N^{l-1} to conclude that $c_1 = 0$ and continue this way to obtain that all the $c_i = 0$.

Next consider the claim that λ^m is the minimal polynomial. If $p\left(\lambda\right)$ is the minimal polynomial, then by the division algorithm,

$$\lambda^m = p\left(\lambda\right) l\left(\lambda\right) + r\left(\lambda\right)$$

where the degree of $r\left(\lambda\right)$ is less than that of $p\left(\lambda\right)$ or else $r\left(\lambda\right) = 0$. The above implies $0 = 0 + r\left(N\right)$ contrary to $p\left(\lambda\right)$ being minimal. Hence $r\left(\lambda\right) = 0$ and so $p\left(\lambda\right)$ divides λ^m. Hence $p\left(\lambda\right) = \lambda^k$ for $k \leq m$. But if $k < m$, this would contradict the definition of m as being the smallest such that $N^m = 0$. ■

For such a nilpotent transformation, let $\left\{\beta_{x_1}, \cdots, \beta_{x_q}\right\}$ be a basis for $\ker\left(N^m\right) = V$ where these β_{x_i} are cyclic. This basis exists thanks to Theorem 9.3.5. Note that you can have $\left|\beta_x\right| < m$ because it is possible for $N^k x = 0$ without $N^k = 0$. Thus

$$V = \text{span}\left(\beta_{x_1}\right) \oplus \cdots \oplus \text{span}\left(\beta_{x_q}\right),$$

each of these subspaces in the above direct sum being N invariant. For x one of the x_k, consider β_x given by

$$x, Nx, N^2 x, \cdots, N^{r-1} x$$

where $N^r x$ is in the span of the above vectors. Then by the above lemma, $N^r x = 0$.

By Theorem 9.2.5, the matrix of N with respect to the above basis is the block diagonal matrix

$$\begin{pmatrix} M^1 & & 0 \\ & \ddots & \\ 0 & & M^q \end{pmatrix}$$

where M^k denotes the matrix of N restricted to span (β_{x_k}). In computing this matrix, I will order β_{x_k} as follows:

$$\left(N^{r_k-1}x_k, \cdots, x_k\right)$$

Also the cyclic sets $\beta_{x_1}, \beta_{x_2}, \cdots, \beta_{x_q}$ will be ordered according to length, the length of β_{x_i} being at least as large as the length of $\beta_{x_{i+1}}, \left|\beta_{x_k}\right| \equiv r_k$. Then since $N^{r_k}x_k = 0$, it is now easy to find M^k. Using the procedure mentioned above for determining the matrix of a linear transformation,

$$\begin{pmatrix} 0 & N^{r_k-1}x_k & \cdots & Nx_k \end{pmatrix} =$$

$$\begin{pmatrix} N^{r_k-1}x_k & N^{r_k-2}x_k & \cdots & x_k \end{pmatrix} \begin{pmatrix} 0 & 1 & & 0 \\ 0 & 0 & \ddots & \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Thus the matrix M_k is the $r_k \times r_k$ matrix which has ones down the super diagonal and zeros elsewhere. The following convenient notation will be used.

Definition 9.4.3 *$J_k(\alpha)$ is a Jordan block if it is a $k \times k$ matrix of the form*

$$J_k(\alpha) = \begin{pmatrix} \alpha & 1 & & 0 \\ 0 & \ddots & \ddots & \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \alpha \end{pmatrix}$$

In words, there is an unbroken string of ones down the super diagonal and the number α filling every space on the main diagonal with zeros everywhere else.

Then with this definition and the above discussion, the following proposition has been proved.

Proposition 9.4.4 *Let $N \in \mathcal{L}(W, W)$ be nilpotent,*

$$N^m = 0$$

for some $m \in \mathbb{N}$. Here W is a p dimensional vector space with field of scalars $\mathbb{F}$. Then there exists a basis for W such that the matrix of N with respect to this basis is of the form

$$J = \begin{pmatrix} J_{r_1}(0) & & & 0 \\ & J_{r_2}(0) & & \\ & & \ddots & \\ 0 & & & J_{r_s}(0) \end{pmatrix} \tag{9.5}$$

where $r_1 \geq r_2 \geq \cdots \geq r_s \geq 1$ and $\sum_{i=1}^{s} r_i = p$. In the above, the $J_{r_j}(0)$ is called a Jordan block of size $r_j \times r_j$ with 0 down the main diagonal.

Observation 9.4.5 *Observe that $J_r(0)^r = 0$ but $J_r(0)^{r-1} \neq 0$.*

In fact, the matrix of the above proposition is unique.

Corollary 9.4.6 *Let J, J' both be matrices of the nilpotent linear transformation $N \in \mathcal{L}(W, W)$ which are of the form described in Proposition 9.4.4. Then $J = J'$. In fact, if the rank of J^k equals the rank of J'^k for all nonnegative integers k, then $J = J'$.*

Proof: Since J and J' are similar, it follows that for each k an integer, J^k and J'^k are similar. Hence, for each k, these matrices have the same rank. Now suppose $J \neq J'$. Note first that

$$J_r(0)^r = 0, \; J_r(0)^{r-1} \neq 0.$$

Denote the blocks of J as $J_{r_k}(0)$ and the blocks of J' as $J_{r'_k}(0)$. Let k be the first such that $J_{r_k}(0) \neq J_{r'_k}(0)$. Suppose that $r_k > r'_k$. By block multiplication and the above observation, it follows that the two matrices J^{r_k-1} and J'^{r_k-1} are respectively of the forms

$$\begin{pmatrix} M_{r_1} & & & & & 0 \\ & \ddots & & & & \\ & & M_{r_k} & & & \\ & & & * & & \\ & & & & \ddots & \\ 0 & & & & & * \end{pmatrix}, \begin{pmatrix} M_{r'_1} & & & & & 0 \\ & \ddots & & & & \\ & & M_{r'_k} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

where $M_{r_j} = M_{r'_j}$ for $j \leq k-1$ but $M_{r'_k}$ is a zero $r'_k \times r'_k$ matrix while M_{r_k} is a larger matrix which is not equal to 0. For example, M_{r_k} could look like

$$M_{r_k} = \begin{pmatrix} 0 & \cdots & 1 \\ & \ddots & \vdots \\ 0 & & 0 \end{pmatrix}$$

Thus there are more pivot columns in J^{r_k-1} than in $(J')^{r_k-1}$, contradicting the requirement that J^k and J'^k have the same rank. ■

9.5 The Jordan Canonical Form

The Jordan canonical form has to do with the case where the minimal polynomial of $A \in \mathcal{L}(V,V)$ splits. Thus there exist λ_k in the field of scalars such that the minimal polynomial of A is of the form

$$p(\lambda) = \prod_{k=1} (\lambda - \lambda_k)^{m_k}$$

Recall the following which follows from Theorem 8.4.4.

Proposition 9.5.1 *Let the minimal polynomial of $A \in \mathcal{L}(V, V)$ be given by*

$$p(\lambda) = \prod_{k=1}^{r} (\lambda - \lambda_k)^{m_k}$$

Then the eigenvalues of A are $\{\lambda_1, \cdots, \lambda_r\}$.

It follows from Corollary 9.2.3 that

$$\begin{aligned} V &= \ker (A - \lambda_1 I)^{m_1} \oplus \cdots \oplus \ker (A - \lambda_r I)^{m_r} \\ &\equiv V_1 \oplus \cdots \oplus V_r \end{aligned}$$

where I denotes the identity linear transformation. Without loss of generality, let the dimensions of the V_k be decreasing from left to right. These V_k are called the generalized eigenspaces.

It follows from the definition of V_k that $(A - \lambda_k I)$ is nilpotent on V_k and clearly each V_k is A invariant. Therefore from Proposition 9.4.4, and letting A_k denote the restriction of A to V_k, there exists an ordered basis for V_k, β_k such that with respect to this basis, the matrix of $(A_k - \lambda_k I)$ is of the form given in that proposition, denoted here by J^k. What is the matrix of A_k with respect to β_k? Letting $\{b_1, \cdots, b_r\} = \beta_k$,

$$A_k b_j = (A_k - \lambda_k I) b_j + \lambda_k I b_j \equiv \sum_s J^k_{sj} b_s + \sum_s \lambda_k \delta_{sj} b_s = \sum_s \left(J^k_{sj} + \lambda_k \delta_{sj}\right) b_s$$

and so the matrix of A_k with respect to this basis is $J^k + \lambda_k I$ where I is the identity matrix.

Therefore, with respect to the ordered basis $\{\beta_1, \cdots, \beta_r\}$ the matrix of A is in Jordan canonical form. This means the matrix is of the form

$$\begin{pmatrix} J(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_r) \end{pmatrix} \tag{9.6}$$

where $J(\lambda_k)$ is an $m_k \times m_k$ matrix of the form

$$\begin{pmatrix} J_{k_1}(\lambda_k) & & & 0 \\ & J_{k_2}(\lambda_k) & & \\ & & \ddots & \\ 0 & & & J_{k_r}(\lambda_k) \end{pmatrix} \tag{9.7}$$

where $k_1 \geq k_2 \geq \cdots \geq k_r \geq 1$ and $\sum_{i=1}^{r} k_i = m_k$. Here $J_k(\lambda)$ is a $k \times k$ Jordan block of the form

$$\begin{pmatrix} \lambda & 1 & & 0 \\ 0 & \lambda & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 0 & \lambda \end{pmatrix} \tag{9.8}$$

This proves the existence part of the following fundamental theorem.

Note that if any of the β_k consists of eigenvectors, then the corresponding Jordan block will consist of a diagonal matrix having λ_k down the main diagonal. This corresponds to $m_k = 1$. The vectors which are in $\ker (A - \lambda_k I)^{m_k}$ which are not in $\ker (A - \lambda_k I)$ are called generalized eigenvectors.

The following is the main result on the Jordan canonical form.

Theorem 9.5.2 *Let V be an n dimensional vector space with field of scalars $\mathbb{C}$ or some other field such that the minimal polynomial of $A \in \mathcal{L}(V,V)$ completely factors into powers of linear factors. Then there exists a unique Jordan canonical form for A as described in 9.6 - 9.8, where uniqueness is in the sense that any two have the same number and size of Jordan blocks.*

Proof: It only remains to verify uniqueness. Suppose there are two, J and J'. Then these are matrices of A with respect to possibly different bases and so they are similar. Therefore, they have the same minimal polynomials and the generalized eigenspaces have the same dimension. Thus the size of the matrices $J(\lambda_k)$ and $J'(\lambda_k)$ defined by the dimension of these generalized eigenspaces, also corresponding to the algebraic multiplicity of λ_k, must be the same. Therefore, they comprise the same set of positive integers. Thus listing the eigenvalues in the same order, corresponding blocks $J(\lambda_k), J'(\lambda_k)$ are the same size.

It remains to show that $J(\lambda_k)$ and $J'(\lambda_k)$ are not just the same size but also are the same up to order of the Jordan blocks running down their respective diagonals. It is only necessary to worry about the number and size of the Jordan blocks making up $J(\lambda_k)$ and $J'(\lambda_k)$. Since J, J' are similar, so are $J - \lambda_k I$ and $J' - \lambda_k I$.

Thus the following two matrices are similar

$$A \equiv \begin{pmatrix} J(\lambda_1) - \lambda_k I & & & & 0 \\ & \ddots & & & \\ & & J(\lambda_k) - \lambda_k I & & \\ & & & \ddots & \\ 0 & & & & J(\lambda_r) - \lambda_k I \end{pmatrix}$$

$$B \equiv \begin{pmatrix} J'(\lambda_1) - \lambda_k I & & & & 0 \\ & \ddots & & & \\ & & J'(\lambda_k) - \lambda_k I & & \\ & & & \ddots & \\ 0 & & & & J'(\lambda_r) - \lambda_k I \end{pmatrix}$$

and consequently, $\text{rank}\left(A^k\right) = \text{rank}\left(B^k\right)$ for all $k \in \mathbb{N}$. Also, both $J(\lambda_j) - \lambda_k I$ and $J'(\lambda_j) - \lambda_k I$ are one to one for every $\lambda_j \neq \lambda_k$. Since all the blocks in both of these matrices are one to one except the blocks $J'(\lambda_k) - \lambda_k I$, $J(\lambda_k) - \lambda_k I$, it follows that this requires the two sequences of numbers $\{\text{rank}\left((J(\lambda_k) - \lambda_k I)^m\right)\}_{m=1}^{\infty}$ and $\left\{\text{rank}\left((J'(\lambda_k) - \lambda_k I)^m\right)\right\}_{m=1}^{\infty}$ must be the same.

Then

$$J(\lambda_k) - \lambda_k I \equiv \begin{pmatrix} J_{k_1}(0) & & & 0 \\ & J_{k_2}(0) & & \\ & & \ddots & \\ 0 & & & J_{k_r}(0) \end{pmatrix}$$

and a similar formula holds for $J'(\lambda_k)$

$$J'(\lambda_k) - \lambda_k I \equiv \begin{pmatrix} J_{l_1}(0) & & & 0 \\ & J_{l_2}(0) & & \\ & & \ddots & \\ 0 & & & J_{l_p}(0) \end{pmatrix}$$

and it is required to verify that $p = r$ and that the same blocks occur in both. Without loss of generality, let the blocks be arranged according to size with the largest on upper left corner falling to smallest in lower right. Now the desired conclusion follows from Corollary 9.4.6. ■

Note that if any of the generalized eigenspaces $\ker (A - \lambda_k I)^{m_k}$ has a basis of eigenvectors, then it would be possible to use this basis and obtain a diagonal matrix in the block corresponding to λ_k. By uniqueness, this is **the** block corresponding to the eigenvalue λ_k. Thus when this happens, the block in the Jordan canonical form corresponding to λ_k is just the diagonal matrix having λ_k down the diagonal and there are **no generalized eigenvectors**.

The Jordan canonical form is very significant when you try to understand powers of a matrix. There exists an $n \times n$ matrix S[1] such that

$$A = S^{-1}JS.$$

Therefore, $A^2 = S^{-1}JSS^{-1}JS = S^{-1}J^2S$ and continuing this way, it follows

$$A^k = S^{-1}J^kS.$$

where J is given in the above corollary. Consider J^k. By block multiplication,

$$J^k = \begin{pmatrix} J_1^k & & 0 \\ & \ddots & \\ 0 & & J_r^k \end{pmatrix}.$$

The matrix J_s is an $m_s \times m_s$ matrix which is of the form

$$J_s = D + N$$

[1]The S here is written as S^{-1} in the corollary.

for D a multiple of the identity and N an upper triangular matrix with zeros down the main diagonal. Thus $N^{m_s} = 0$. Now since D is just a multiple of the identity, it follows that $DN = ND$. Therefore, the usual binomial theorem may be applied and this yields the following equations for $k \geq m_s$.

$$\begin{aligned} J_s^k &= (D+N)^k = \sum_{j=0}^{k} \binom{k}{j} D^{k-j} N^j \\ &= \sum_{j=0}^{m_s} \binom{k}{j} D^{k-j} N^j, \end{aligned} \tag{9.9}$$

the third equation holding because $N^{m_s} = 0$. Thus J_s^k is of the form

$$J_s^k = \begin{pmatrix} \alpha^k & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha^k \end{pmatrix}.$$

Lemma 9.5.3 *Suppose J is of the form J_s, a Jordan block where the constant α, on the main diagonal is less than one in absolute value. Then*

$$\lim_{k \to \infty} \left(J^k\right)_{ij} = 0.$$

Proof: From 9.9, it follows that for large k, and $j \leq m_s$,

$$\binom{k}{j} \leq \frac{k(k-1)\cdots(k-m_s+1)}{m_s!}.$$

Therefore, letting C be the largest value of $\left|\left(N^j\right)_{pq}\right|$ for $0 \leq j \leq m_s$,

$$\left|\left(J^k\right)_{pq}\right| \leq m_s C \left(\frac{k(k-1)\cdots(k-m_s+1)}{m_s!}\right) |\alpha|^{k-m_s}$$

which converges to zero as $k \to \infty$. This is most easily seen by applying the ratio test to the series

$$\sum_{k=m_s}^{\infty} \left(\frac{k(k-1)\cdots(k-m_s+1)}{m_s!}\right) |\alpha|^{k-m_s}$$

and then noting that if a series converges, then the k^{th} term converges to zero. ■

9.6 Exercises

1. In the discussion of Nilpotent transformations, it was asserted that if two $n \times n$ matrices A, B are similar, then A^k is also similar to B^k. Why is this so? If two matrices are similar, why must they have the same rank?

2. If A, B are both invertible, then they are both row equivalent to the identity matrix. Are they necessarily similar? Explain.

3. Suppose you have two nilpotent matrices A, B and A^k and B^k both have the same rank for all $k \geq 1$. Does it follow that A, B are similar? What if it is not known that A, B are nilpotent? Does it follow then?

4. When we say a polynomial equals zero, we mean that all the coefficients equal 0. If we assign a different meaning to it which says that a polynomial $p(\lambda)$ equals zero when it is the zero function, ($p(\lambda) = 0$ for every $\lambda \in \mathbb{F}$.) does this amount to the same thing? Is there any difference in the two definitions for ordinary fields like $\mathbb{Q}$? **Hint:** Consider for the field of scalars $\mathbb{Z}_2$, the integers mod 2 and consider $p(\lambda) = \lambda^2 + \lambda$.

5. Let $A \in \mathcal{L}(V,V)$ where V is a finite dimensional vector space with field of scalars $\mathbb{F}$. Let $p(\lambda)$ be the minimal polynomial and suppose $\phi(\lambda)$ is any nonzero polynomial such that $\phi(A)$ is not one to one and $\phi(\lambda)$ has smallest possible degree such that $\phi(A)$ is nonzero and not one to one. Show $\phi(\lambda)$ must divide $p(\lambda)$.

6. Let $A \in \mathcal{L}(V,V)$ where V is a finite dimensional vector space with field of scalars $\mathbb{F}$. Let $p(\lambda)$ be the minimal polynomial and suppose $\phi(\lambda)$ is an irreducible polynomial with the property that $\phi(A)x = 0$ for some specific $x \neq 0$. Show that $\phi(\lambda)$ must divide $p(\lambda)$. **Hint:** First write $p(\lambda) = \phi(\lambda) g(\lambda) + r(\lambda)$ where $r(\lambda)$ is either 0 or has degree smaller than the degree of $\phi(\lambda)$. If $r(\lambda) = 0$ you are done. Suppose it is not 0. Let $\eta(\lambda)$ be the monic polynomial of smallest degree with the property that $\eta(A)x = 0$. Now use the Euclidean algorithm to divide $\phi(\lambda)$ by $\eta(\lambda)$. Contradict the irreducibility of $\phi(\lambda)$.

7. Suppose A is a linear transformation and let the characteristic polynomial be

$$\det(\lambda I - A) = \prod_{j=1}^{q} \phi_j(\lambda)^{n_j}$$

where the $\phi_j(\lambda)$ are irreducible. Explain using Corollary 7.3.11 why the irreducible factors of the minimal polynomial are $\phi_j(\lambda)$ and why the minimal polynomial is of the form $\prod_{j=1}^{q} \phi_j(\lambda)^{r_j}$ where $r_j \leq n_j$. You can use the Cayley Hamilton theorem if you like.

8. Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Find the minimal polynomial for A.

9. Suppose A is an $n \times n$ matrix and let $\mathbf{v}$ be a vector. Consider the A cyclic set of vectors $\{\mathbf{v}, A\mathbf{v}, \cdots, A^{m-1}\mathbf{v}\}$ where this is an independent set of vectors but $A^m\mathbf{v}$ is a linear combination of the preceding vectors in the list. Show how to obtain a monic polynomial of smallest degree, m, $\phi_{\mathbf{v}}(\lambda)$ such that

$$\phi_{\mathbf{v}}(A)\mathbf{v} = \mathbf{0}$$

Now let $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ be a basis and let $\phi(\lambda)$ be the least common multiple of the $\phi_{\mathbf{w}_k}(\lambda)$. Explain why this must be the minimal polynomial of A. Give a reasonably easy algorithm for computing $\phi_{\mathbf{v}}(\lambda)$.

10. Here is a matrix.

$$\begin{pmatrix} -7 & -1 & -1 \\ -21 & -3 & -3 \\ 70 & 10 & 10 \end{pmatrix}$$

Using the process of Problem 9 find the minimal polynomial of this matrix. It turns out the characteristic polynomial is λ^3.

11. Find the minimal polynomial for

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ -3 & 2 & 1 \end{pmatrix}$$

by the above technique. Is what you found also the characteristic polynomial?

12. Let A be an $n \times n$ matrix with field of scalars $\mathbb{C}$. Letting λ be an eigenvalue, show the dimension of the eigenspace equals the number of Jordan blocks in the Jordan canonical form which are associated with λ. Recall the eigenspace is $\ker(\lambda I - A)$.

13. For any $n \times n$ matrix, why is the dimension of the eigenspace always less than or equal to the algebraic multiplicity of the eigenvalue as a root of the characteristic equation? **Hint:** Note the algebraic multiplicity is the size of the appropriate block in the Jordan form.

14. Give an example of two nilpotent matrices which are not similar but have the same minimal polynomial if possible.

15. Use the existence of the Jordan canonical form for a linear transformation whose minimal polynomial factors completely to give a proof of the Cayley Hamilton theorem which is valid for any field of scalars. **Hint:** First assume the minimal polynomial factors completely into linear factors. If this does not happen, consider a splitting field of the minimal polynomial. Then consider the minimal polynomial with respect to this larger field. How will the two minimal polynomials be related? Show the minimal polynomial always divides the characteristic polynomial.

16. Here is a matrix. Find its Jordan canonical form by directly finding the eigenvectors and generalized eigenvectors based on these to find a basis which will yield the Jordan form. The eigenvalues are 1 and 2.

$$\begin{pmatrix} -3 & -2 & 5 & 3 \\ -1 & 0 & 1 & 2 \\ -4 & -3 & 6 & 4 \\ -1 & -1 & 1 & 3 \end{pmatrix}$$

Why is it typically impossible to find the Jordan canonical form?

17. People like to consider the solutions of first order linear systems of equations which are of the form

$$\mathbf{x}'(t) = A\mathbf{x}(t)$$

where here A is an $n \times n$ matrix. From the theorem on the Jordan canonical form, there exist S and S^{-1} such that $A = SJS^{-1}$ where J is a Jordan form. Define $\mathbf{y}(t) \equiv S^{-1}\mathbf{x}(t)$. Show $\mathbf{y}' = J\mathbf{y}$. Now suppose $\Psi(t)$ is an $n \times n$ matrix whose columns are solutions of the above differential equation. Thus

$$\Psi' = A\Psi$$

Now let Φ be defined by $S\Phi S^{-1} = \Psi$. Show

$$\Phi' = J\Phi.$$

18. In the above Problem show that

$$\det(\Psi)' = \operatorname{trace}(A)\det(\Psi)$$

and so

$$\det(\Psi(t)) = Ce^{\operatorname{trace}(A)t}$$

This is called Abel's formula and $\det(\Psi(t))$ is called the Wronskian. **Hint:** Show it suffices to consider

$$\Phi' = J\Phi$$

and establish the formula for Φ. Next let

$$\Phi = \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_n \end{pmatrix}$$

where the ϕ_j are the rows of Φ. Then explain why

$$\det(\Phi)' = \sum_{i=1}^{n} \det(\Phi_i) \tag{9.10}$$

where Φ_i is the same as Φ except the i^{th} row is replaced with ϕ_i' instead of the row ϕ_i. Now from the form of J,

$$\Phi' = D\Phi + N\Phi$$

where N has all nonzero entries above the main diagonal. Explain why

$$\phi_i'(t) = \lambda_i\phi_i(t) + a_i\phi_{i+1}(t)$$

Now use this in the formula for the derivative of the Wronskian given in 9.10 and use properties of determinants to obtain

$$\det(\Phi)' = \sum_{i=1}^{n} \lambda_i \det(\Phi).$$

Obtain Abel's formula

$$\det(\Phi) = Ce^{\operatorname{trace}(A)t}$$

and so the Wronskian $\det\Phi$ either vanishes identically or never.

19. Let A be an $n \times n$ matrix and let J be its Jordan canonical form. Recall J is a block diagonal matrix having blocks $J_k(\lambda)$ down the diagonal. Each of these blocks is of the form

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

Now for $\varepsilon > 0$ given, let the diagonal matrix D_ε be given by

$$D_\varepsilon = \begin{pmatrix} 1 & & & 0 \\ & \varepsilon & & \\ & & \ddots & \\ 0 & & & \varepsilon^{k-1} \end{pmatrix}$$

Show that $D_\varepsilon^{-1} J_k(\lambda) D_\varepsilon$ has the same form as $J_k(\lambda)$ but instead of ones down the super diagonal, there is ε down the super diagonal. That is $J_k(\lambda)$ is replaced with

$$\begin{pmatrix} \lambda & \varepsilon & & 0 \\ & \lambda & \ddots & \\ & & \ddots & \varepsilon \\ 0 & & & \lambda \end{pmatrix}$$

Now show that for A an $n \times n$ matrix, it is similar to one which is just like the Jordan canonical form except instead of the blocks having 1 down the super diagonal, it has ε.

20. Let A be in $\mathcal{L}(V, V)$ and suppose that $A^p x \neq 0$ for some $x \neq 0$. Show that $A^p e_k \neq 0$ for some $e_k \in \{e_1, \cdots, e_n\}$, a basis for V. If you have a matrix which is nilpotent, ($A^m = 0$ for some m) will it always be possible to find its Jordan form? Describe how to do it if this is the case. **Hint:** First explain why all the eigenvalues are 0. Then consider the way the Jordan form for nilpotent transformations was constructed in the above.

21. Suppose A is an $n \times n$ matrix and that it has n distinct eigenvalues. How do the minimal polynomial and characteristic polynomials compare? Determine other conditions based on the Jordan Canonical form which will cause the minimal and characteristic polynomials to be different.

22. Suppose A is a 3×3 matrix and it has at least two distinct eigenvalues. Is it possible that the minimal polynomial is different than the characteristic polynomial?

23. If A is an $n \times n$ matrix of entries from a field of scalars and if the minimal polynomial of A splits over this field of scalars, does it follow that the characteristic polynomial of A also splits? Explain why or why not.

24. Show that if two $n \times n$ matrices A, B are similar, then they have the same minimal polynomial and also that if this minimal polynomial is of the form $p(\lambda) = \prod_{i=1}^{s} \phi_i(\lambda)^{r_i}$ where the $\phi_i(\lambda)$ are irreducible and monic, then $\ker(\phi_i(A)^{r_i})$ and $\ker(\phi_i(B)^{r_i})$ have the same dimension. Why is this so? This was what was responsible for the blocks corresponding to an eigenvalue being of the same size.

25. Show that a given complex $n \times n$ matrix is non defective (diagonalizable) if and only if the minimal polynomial has no repeated roots.

26. Describe a straight forward way to determine the minimal polynomial of an $n \times n$ matrix using row operations. Next show that if $p(\lambda)$ and $p'(\lambda)$ are relatively prime, then $p(\lambda)$ has no repeated roots. With the above problem, explain how this gives a way to determine whether a matrix is non defective.

27. In Theorem 9.3.5 show that each cyclic set β_x is associated with a monic polynomial $\eta_x(\lambda)$ such that $\eta_x(A)(x) = 0$ and this polynomial has smallest possible degree such that this happens. Show that the cyclic sets β_{x_i} can be arranged such that $\eta_{x_{i+1}}(\lambda) / \eta_{x_i}(\lambda)$.

28. Show that if A is a complex $n \times n$ matrix, then A and A^T are similar. **Hint:** Consider a Jordan block. Note that

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}$$

29. Let A be a linear transformation defined on a finite dimensional vector space V. Let the minimal polynomial be $\prod_{i=1}^{q} \phi_i(\lambda)^{m_i}$ and let $\left(\beta_{v_1^i}^i, \cdots, \beta_{v_{r_i}^i}^i\right)$ be the cyclic sets such that $\left\{\beta_{v_1^i}^i, \cdots, \beta_{v_{r_i}^i}^i\right\}$ is a basis for $\ker\left(\phi_i(A)^{m_i}\right)$. Let $v = \sum_i \sum_j v_j^i$. Now let $q(\lambda)$ be any polynomial and suppose that

$$q(A)\, v = 0$$

Show that it follows $q(A) = 0$. **Hint:** First consider the special case where a basis for V is $\left\{x, Ax, \cdots, A^{n-1}x\right\}$ and $q(A)\, x = 0$.

9.7 The Rational Canonical Form*

Here one has the minimal polynomial in the form $\prod_{k=1}^{q} \phi(\lambda)^{m_k}$ where $\phi(\lambda)$ is an irreducible monic polynomial. It is not necessarily the case that $\phi(\lambda)$ is a linear factor. Thus this case is completely general and includes the situation where the field is arbitrary. In particular, it includes the case where the field of scalars is, for example, the rational numbers. This may be partly why it is called the rational canonical form. As you know, the rational numbers are notorious for not having roots to polynomial equations which have integer or rational coefficients.

This canonical form is due to Frobenius. I am following the presentation given in [10] and there are more details given in this reference. Another good source which has additional results is [15].

Here is a definition of the concept of a companion matrix.

Definition 9.7.1 *Let*

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

be a monic polynomial. The companion matrix of $q(\lambda)$, *denoted as* $C(q(\lambda))$ *is the matrix*

$$\begin{pmatrix} 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & -a_1 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & -a_{n-1} \end{pmatrix}$$

Proposition 9.7.2 *Let* $q(\lambda)$ *be a polynomial and let* $C(q(\lambda))$ *be its companion matrix. Then* $q(C(q(\lambda))) = 0$.

Proof: Write C instead of $C(q(\lambda))$ for short. Note that

$$C\mathbf{e}_1 = \mathbf{e}_2, C\mathbf{e}_2 = \mathbf{e}_3, \cdots, C\mathbf{e}_{n-1} = \mathbf{e}_n$$

Thus

$$\mathbf{e}_k = C^{k-1}\mathbf{e}_1,\ k = 1, \cdots, n \tag{9.11}$$

and so it follows

$$\left\{\mathbf{e}_1, C\mathbf{e}_1, C^2\mathbf{e}_1, \cdots, C^{n-1}\mathbf{e}_1\right\} \tag{9.12}$$

are linearly independent. Hence these form a basis for $\mathbb{F}^n$. Now note that $C\mathbf{e}_n$ is given by

$$C\mathbf{e}_n = -a_0\mathbf{e}_1 - a_1\mathbf{e}_2 - \cdots - \mathbf{a}_{n-1}\mathbf{e}_n$$

and from 9.11 this implies

$$C^n\mathbf{e}_1 = -a_0\mathbf{e}_1 - a_1 C\mathbf{e}_1 - \cdots - \mathbf{a}_{n-1}C^{n-1}\mathbf{e}_1$$

and so $q(C)\mathbf{e}_1 = \mathbf{0}$. Now since 9.12 is a basis, every vector of $\mathbb{F}^n$ is of the form $k(C)\mathbf{e}_1$ for some polynomial $k(\lambda)$. Therefore, if $\mathbf{v} \in \mathbb{F}^n$,

$$q(C)\mathbf{v} = q(C)k(C)\mathbf{e}_1 = k(C)q(C)\mathbf{e}_1 = \mathbf{0}$$

which shows $q(C) = 0$. ■

The following theorem is on the existence of the rational canonical form.

Theorem 9.7.3 *Let* $A \in \mathcal{L}(V, V)$ *where* V *is a vector space with field of scalars* $\mathbb{F}$ *and minimal polynomial* $\prod_{i=1}^{q} \phi_i(\lambda)^{m_i}$ *where each* $\phi_i(\lambda)$ *is irreducible and monic. Letting* $V_k \equiv \ker\left(\phi_k(\lambda)^{m_k}\right)$, *it follows*

$$V = V_1 \oplus \cdots \oplus V_q$$

where each V_k *is* A *invariant. Letting* B_k *denote a basis for* V_k *and* M^k *the matrix of the restriction of* A *to* V_k, *it follows that the matrix of* A *with respect to the basis* $\{B_1, \cdots, B_q\}$ *is the block diagonal matrix of the form*

$$\begin{pmatrix} M^1 & & 0 \\ & \ddots & \\ 0 & & M^q \end{pmatrix} \tag{9.13}$$

If B_k *is given as* $\left\{\beta_{v_1}, \cdots, \beta_{v_s}\right\}$ *as described in Theorem 9.3.5 where each* β_{v_j} *is an* A *cyclic set of vectors, then the matrix* M^k *is of the form*

$$M^k = \begin{pmatrix} C\left(\phi_k(\lambda)^{r_1}\right) & & 0 \\ & \ddots & \\ 0 & & C\left(\phi_k(\lambda)^{r_s}\right) \end{pmatrix} \tag{9.14}$$

where the A *cyclic sets of vectors may be arranged in order such that the positive integers* r_j *satisfy* $r_1 \geq \cdots \geq r_s$ *and* $C\left(\phi_k(\lambda)^{r_j}\right)$ *is the companion matrix of the polynomial* $\phi_k(\lambda)^{r_j}$.

Proof: By Theorem 9.2.5 the matrix of A with respect to $\{B_1, \cdots, B_q\}$ is of the form given in 9.13. Now by Theorem 9.3.5 the basis B_k may be chosen in the form $\{\beta_{v_1}, \cdots, \beta_{v_s}\}$ where each β_{v_k} is an A cyclic set of vectors and also it can be assumed the lengths of these β_{v_k} are decreasing. Thus

$$V_k = \text{span}\left(\beta_{v_1}\right) \oplus \cdots \oplus \text{span}\left(\beta_{v_s}\right)$$

and it only remains to consider the matrix of A restricted to span $\left(\beta_{v_k}\right)$. Then you can apply Theorem 9.2.5 to get the result in 9.14. Say

$$\beta_{v_k} = v_k, Av_k, \cdots, A^{d-1}v_k$$

where $\eta(A) v_k = 0$ and the degree of $\eta(\lambda)$ is d, the smallest degree such that this is so, η being a monic polynomial. Then $\eta(\lambda)$ must divide $\phi_k(\lambda)^{m_k}$. By Corollary 7.3.11, $\eta(\lambda) = \phi_k(\lambda)^{r_k}$ where $r_k \leq m_k$. It remains to consider the matrix of A restricted to span $\left(\beta_{v_k}\right)$. Say

$$\eta(\lambda) = \phi_k(\lambda)^{r_k} = a_0 + a_1\lambda + \cdots + a_{d-1}\lambda^{d-1} + \lambda^d$$

Thus, since $\eta(A) v_k = 0$,

$$A^d v_k = -a_0 v_k - a_1 A v_k - \cdots - a_{d-1} A^{d-1} v_k$$

Recall the formalism for finding the matrix of A restricted to this invariant subspace.

$$\begin{pmatrix} Av_k & A^2v_k & A^3v_k & \cdots & -a_0v_k - a_1Av_k - \cdots - a_{d-1}A^{d-1}v_k \end{pmatrix} =$$

$$\begin{pmatrix} v_k & Av_k & A^2v_k & \cdots & A^{d-1}v_k \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & \cdots & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & \ddots & & \vdots \\ & \ddots & \ddots & 0 & -a_{d-2} \\ 0 & & 0 & 1 & -a_{d-1} \end{pmatrix}$$

Thus the matrix of the transformation is the above. This is the companion matrix of $\phi_k(\lambda)^{r_k} = \eta(\lambda)$. In other words, $C = C\left(\phi_k(\lambda)^{r_k}\right)$ and so M^k has the form claimed in the theorem. ■

9.8 Uniqueness

Given $A \in \mathcal{L}(V, V)$ where V is a vector space having field of scalars $\mathbb{F}$, the above shows there exists a rational canonical form for A. Could A have more than one rational canonical form? Recall the definition of an A cyclic set. For convenience, here it is again.

Definition 9.8.1 *Letting $x \neq 0$ denote by β_x the vectors $\{x, Ax, A^2x, \cdots, A^{m-1}x\}$ where m is the smallest such that $A^m x \in \text{span}\left(x, \cdots, A^{m-1}x\right)$.*

The following proposition ties these A cyclic sets to polynomials. It is just a review of ideas used above to prove existence.

Proposition 9.8.2 *Let $x \neq 0$ and consider $\{x, Ax, A^2x, \cdots, A^{m-1}x\}$. Then this is an A cyclic set if and only if there exists a monic polynomial $\eta(\lambda)$ such that $\eta(A) x = 0$ and among all such polynomials $\psi(\lambda)$ satisfying $\psi(A) x = 0$, $\eta(\lambda)$ has the smallest degree. If $V = \ker\left(\phi(\lambda)^m\right)$ where $\phi(\lambda)$ is monic and irreducible, then for some positive integer $p \leq m, \eta(\lambda) = \phi(\lambda)^p$.*

The following is the main consideration for proving uniqueness. It will depend on what was already shown for the Jordan canonical form. This will apply to the nilpotent matrix $\phi(A)$.

Lemma 9.8.3 *Let V be a vector space and $A \in \mathcal{L}(V,V)$ has minimal polynomial $\phi(\lambda)^m$ where $\phi(\lambda)$ is irreducible and has degree d. Let the basis for V consist of $\{\beta_{v_1}, \cdots, \beta_{v_s}\}$ where β_{v_k} is A cyclic as described above and the rational canonical form for A is the matrix taken with respect to this basis. Then letting $|\beta_{v_k}|$ denote the number of vectors in β_{v_k}, it follows there is only one possible set of numbers $|\beta_{v_k}|$.*

Proof: Say β_{v_j} is associated with the polynomial $\phi(\lambda)^{p_j}$. Thus, as described above $|\beta_{v_j}|$ equals $p_j d$. Consider the following table which comes from the A cyclic set

$$\{v_j, Av_j, \cdots, A^{d-1}v_j, \cdots, A^{p_j d-1}v_j\}$$

α_0^j	α_1^j	α_2^j	$\cdots$	α_{d-1}^j
v_j	Av_j	$A^2 v_j$	$\cdots$	$A^{d-1}v_j$
$\phi(A) v_j$	$\phi(A) Av_j$	$\phi(A) A^2 v_j$	$\cdots$	$\phi(A) A^{d-1}v_j$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$\phi(A)^{p_j-1} v_j$	$\phi(A)^{p_j-1} Av_j$	$\phi(A)^{p_j-1} A^2 v_j$	$\cdots$	$\phi(A)^{p_j-1} A^{d-1}v_j$

In the above, α_k^j signifies the vectors below it in the k^{th} column. None of these vectors below the top row are equal to 0 because the degree of $\phi(\lambda)^{p_j-1}\lambda^{d-1}$ is $dp_j - 1$, which is less than $p_j d$ and the smallest degree of a nonzero polynomial sending v_j to 0 is $p_j d$. Also, each of these vectors is in the span of β_{v_j} and there are dp_j of them, just as there are dp_j vectors in β_{v_j}.

Claim: The vectors $\left\{\alpha_0^j, \cdots, \alpha_{d-1}^j\right\}$ are linearly independent.

Proof of claim: Suppose

$$\sum_{i=0}^{d-1}\sum_{k=0}^{p_j-1} c_{ik}\phi(A)^k A^i v_j = 0$$

Then multiplying both sides by $\phi(A)^{p_j-1}$ this yields

$$\sum_{i=0}^{d-1} c_{i0}\phi(A)^{p_j-1} A^i v_j = 0$$

this is because if $k \geq 1$, you have a typical term of the form

$$c_{ik}\phi(A)^{p_j-1}\phi(A)^k A^i v_j = A^i\phi(A)^{k-1} c_{ik}\phi(A)^{p_j} v_j = 0$$

Now if any of the c_{i0} is nonzero this would imply there exists a polynomial having degree smaller than $p_j d$ which sends v_j to 0. In fact, the polynomial would have degree $d-1+p_j-1$. Since this does not happen, it follows each $c_{i0} = 0$. Thus

$$\sum_{i=0}^{d-1}\sum_{k=1}^{p_j-1} c_{ik}\phi(A)^k A^i v_j = 0$$

Now multiply both sides by $\phi(A)^{p_j-2}$ and do a similar argument to assert that $c_{i1} = 0$ for each i. Continuing this way, all the $c_{ik} = 0$ and this proves the claim.

Thus the vectors $\left\{\alpha_0^j, \cdots, \alpha_{d-1}^j\right\}$ are linearly independent and there are $p_j d = \left|\beta_{v_j}\right|$ of them. Therefore, they form a basis for span $\left(\beta_{v_j}\right)$. Also note that if you list the columns in reverse order starting from the bottom and going toward the top, the vectors $\left\{\alpha_0^j, \cdots, \alpha_{d-1}^j\right\}$ yield Jordan blocks in the matrix of $\phi(A)$. Hence, considering all these vectors $\left\{\alpha_0^j, \cdots, \alpha_{d-1}^j\right\}_{j=1}^s$, each listed in the reverse order, the matrix of $\phi(A)$ with respect to this basis of V is in Jordan canonical form. See Proposition 9.4.4 and Theorem 9.5.2 on existence and uniqueness for the Jordan form. This Jordan form is unique up to order of the blocks. For a given j $\left\{\alpha_0^j, \cdots, \alpha_{d-1}^j\right\}$ yields d Jordan blocks of size p_j for $\phi(A)$. The size and number of Jordan blocks of $\phi(A)$ depends only on $\phi(A)$, hence only on A. Once A is determined, $\phi(A)$ is determined and hence the number and size of Jordan blocks is determined, so the exponents p_j are determined and this shows the lengths of the $\beta_{v_j}, p_j d$ are also determined. ■

Note that if the p_j are known, then so is the rational canonical form because it comes from blocks which are companion matrices of the polynomials $\phi(\lambda)^{p_j}$. Now here is the main result.

Theorem 9.8.4 *Let V be a vector space having field of scalars $\mathbb{F}$ and let $A \in \mathcal{L}(V,V)$. Then the rational canonical form of A is unique up to order of the blocks.*

Proof: Let the minimal polynomial of A be $\prod_{k=1}^{q}\phi_k(\lambda)^{m_k}$. Then recall from Corollary 9.2.3

$$V = V_1 \oplus \cdots \oplus V_q$$

where $V_k = \ker\left(\phi_k(A)^{m_k}\right)$. Also recall from Corollary 9.2.4 that the minimal polynomial of the restriction of A to V_k is $\phi_k(\lambda)^{m_k}$. Now apply Lemma 9.8.3 to A restricted to V_k. ■

In the case where two $n \times n$ matrices M, N are similar, recall this is equivalent to the two being matrices of the same linear transformation taken with respect to two different bases. Hence each are similar to the same rational canonical form.

Example 9.8.5 *Here is a matrix.*

$$A = \begin{pmatrix} 5 & -2 & 1 \\ 2 & 10 & -2 \\ 9 & 0 & 9 \end{pmatrix}$$

Find a similarity transformation which will produce the rational canonical form for A.

The minimal polynomial is $\lambda^3 - 24\lambda^2 + 180\lambda - 432$. Why? This factors as

$$(\lambda - 6)^2 (\lambda - 12)$$

Thus $\mathbb{Q}^3$ is the direct sum of $\ker\left((A - 6I)^2\right)$ and $\ker(A - 12I)$. Consider the first of these. You see easily that this is

$$y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, y, z \in \mathbb{Q}.$$

What about the length of A cyclic sets? It turns out it doesn't matter much. You can start with either of these and get a cycle of length 2. Lets pick the second one. This leads to the cycle

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -4 \\ -4 \\ 0 \end{pmatrix} = A \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -12 \\ -48 \\ -36 \end{pmatrix} = A^2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

where the last of the three is a linear combination of the first two. Take the first two as the first two columns of S. To get the third, you need a cycle of length 1 corresponding to $\ker(A - 12I)$. This yields the eigenvector $\begin{pmatrix} 1 & -2 & 3 \end{pmatrix}^T$. Thus

$$S = \begin{pmatrix} -1 & -4 & 1 \\ 0 & -4 & -2 \\ 1 & 0 & 3 \end{pmatrix}$$

Now using Proposition 8.3.10, the Rational canonical form for A should be

$$\begin{pmatrix} -1 & -4 & 1 \\ 0 & -4 & -2 \\ 1 & 0 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -2 & 1 \\ 2 & 10 & -2 \\ 9 & 0 & 9 \end{pmatrix} \begin{pmatrix} -1 & -4 & 1 \\ 0 & -4 & -2 \\ 1 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & -36 & 0 \\ 1 & 12 & 0 \\ 0 & 0 & 12 \end{pmatrix}$$

Example 9.8.6 *Here is a matrix.*

$$A = \begin{pmatrix} 12 & -3 & -19 & -14 & 8 \\ -4 & 1 & 1 & 6 & -4 \\ 4 & 5 & 5 & -2 & 4 \\ 0 & -5 & -5 & 2 & 0 \\ -4 & 3 & 11 & 6 & 0 \end{pmatrix}$$

Find a basis such that if S is the matrix which has these vectors as columns $S^{-1}AS$ is in rational canonical form assuming the field of scalars is $\mathbb{Q}$.

First it is necessary to find the minimal polynomial. Of course you can find the characteristic polynomial and then take away factors till you find the minimal polynomial. However, there is a much better way which is described in the exercises. Leaving out this detail, the minimal polynomial is

$$\lambda^3 - 12\lambda^2 + 64\lambda - 128$$

This polynomial factors as

$$(\lambda - 4)\left(\lambda^2 - 8\lambda + 32\right) \equiv \phi_1(\lambda)\,\phi_2(\lambda)$$

where the second factor is irreducible over $\mathbb{Q}$. Consider $\phi_2(\lambda)$ first. Messy computations yield

$$\ker(\phi_2(A)) = a\begin{pmatrix}-1\\1\\0\\0\\0\end{pmatrix} + b\begin{pmatrix}-1\\0\\1\\0\\0\end{pmatrix} + c\begin{pmatrix}-1\\0\\0\\1\\0\end{pmatrix} + d\begin{pmatrix}-2\\0\\0\\0\\1\end{pmatrix}.$$

Now start with one of these basis vectors and look for an A cycle. Picking the first one, you obtain the cycle

$$\begin{pmatrix}-1\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}-15\\5\\1\\-5\\7\end{pmatrix}$$

because the next vector involving A^2 yields a vector which is in the span of the above two. You check this by making the vectors the columns of a matrix and finding the row reduced echelon form. Clearly this cycle does not span $\ker(\phi_2(A))$, so look for another cycle. Begin with a vector which is not in the span of these two. The last one works well. Thus another A cycle is

$$\begin{pmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -16 \\ 4 \\ -4 \\ 0 \\ 8 \end{pmatrix}$$

It follows a basis for $\ker(\phi_2(A))$ is

$$\left\{ \begin{pmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -16 \\ 4 \\ -4 \\ 0 \\ 8 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -15 \\ 5 \\ 1 \\ -5 \\ 7 \end{pmatrix} \right\}$$

Finally consider a cycle coming from $\ker(\phi_1(A))$. This amounts to nothing more than finding an eigenvector for A corresponding to the eigenvalue 4. An eigenvector is

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 1 \end{pmatrix}^T$$

Now the desired matrix for the similarity transformation is

$$S \equiv \begin{pmatrix} -2 & -16 & -1 & -15 & -1 \\ 0 & 4 & 1 & 5 & 0 \\ 0 & -4 & 0 & 1 & 0 \\ 0 & 0 & 0 & -5 & 0 \\ 1 & 8 & 0 & 7 & 1 \end{pmatrix}$$

Then doing the computations, you get

$$S^{-1}AS = \begin{pmatrix} 0 & -32 & 0 & 0 & 0 \\ 1 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & -32 & 0 \\ 0 & 0 & 1 & 8 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

and you see this is in rational canonical form, the two 2×2 blocks being companion matrices for the polynomial $\lambda^2 - 8\lambda + 32$ and the 1×1 block being a companion matrix for $\lambda - 4$. Note that you could have written this without finding a similarity transformation to produce it. This follows from the above theory which gave the existence of the rational canonical form.

Obviously there is a lot more which could be considered about rational canonical forms. Just begin with a strange field and start investigating what can be said. One can also derive more systematic methods for finding the rational canonical form. The advantage of this is you don't need to find the eigenvalues in order to compute the rational canonical form and it can often be computed for this reason, unlike the Jordan form. The uniqueness of this rational canonical form can be used to determine whether two matrices consisting of entries in some field are similar.

9.9 Exercises

1. Suppose A is a linear transformation and let the characteristic polynomial be

$$\det(\lambda I - A) = \prod_{j=1}^{q} \phi_j(\lambda)^{n_j}$$

where the $\phi_j(\lambda)$ are irreducible. Explain using Corollary 7.3.11 why the irreducible factors of the minimal polynomial are $\phi_j(\lambda)$ and why the minimal polynomial is of the form $\prod_{j=1}^{q} \phi_j(\lambda)^{r_j}$ where $r_j \leq n_j$. You can use the Cayley Hamilton theorem if you like.

2. Find the minimal polynomial for

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ -3 & 2 & 1 \end{pmatrix}$$

by the above technique assuming the field of scalars is the rational numbers. Is what you found also the characteristic polynomial?

3. Show, using the rational root theorem, the minimal polynomial for A in the above problem is irreducible with respect to $\mathbb{Q}$. Letting the field of scalars be $\mathbb{Q}$ find the rational canonical form and a similarity transformation which will produce it.

4. Letting the field of scalars be $\mathbb{Q}$, find the rational canonical form for the matrix

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 2 & 3 & 0 & 2 \\ 1 & 3 & 2 & 4 \\ 1 & 2 & 1 & 2 \end{pmatrix}$$

5. Let $A : \mathbb{Q}^3 \to \mathbb{Q}^3$ be linear. Suppose the minimal polynomial is $(\lambda - 2)\left(\lambda^2 + 2\lambda + 7\right)$. Find the rational canonical form. Can you give generalizations of this rather simple problem to other situations?

6. Find the rational canonical form with respect to the field of scalars equal to $\mathbb{Q}$ for the matrix

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

Observe that this particular matrix is already a companion matrix of $\lambda^3 - \lambda^2 + \lambda - 1$. Then find the rational canonical form if the field of scalars equals $\mathbb{C}$ or $\mathbb{Q} + i\mathbb{Q}$.

7. Let $q(\lambda)$ be a polynomial and C its companion matrix. Show the characteristic and minimal polynomial of C are the same and both equal $q(\lambda)$.

8. ↑Use the existence of the rational canonical form to give a proof of the Cayley Hamilton theorem valid for any field, even fields like the integers mod p for p a prime. The earlier proof based on determinants was fine for fields like $\mathbb{Q}$ or $\mathbb{R}$ where you could let $\lambda \to \infty$ but it is not clear the same result holds in general.

9. Suppose you have two $n \times n$ matrices A, B whose entries are in a field $\mathbb{F}$ and suppose $\mathbb{G}$ is an extension of $\mathbb{F}$. For example, you could have $\mathbb{F} = \mathbb{Q}$ and $\mathbb{G} = \mathbb{C}$. Suppose A and B are similar with respect to the field $\mathbb{G}$. Can it be concluded that they are similar with respect to the field $\mathbb{F}$? **Hint:** First show that the two have the same minimal polynomial over $\mathbb{F}$. Next consider the proof of Lemma 9.8.3 and show that they have the same rational canonical form with respect to $\mathbb{F}$.

Chapter 10

Markov Processes

10.1 Regular Markov Matrices

The existence of the Jordan form is the basis for the proof of limit theorems for certain kinds of matrices called Markov matrices.

Definition 10.1.1 *An $n \times n$ matrix $A = (a_{ij})$, is a Markov matrix if $a_{ij} \geq 0$ for all i, j and*

$$\sum_i a_{ij} = 1.$$

It may also be called a stochastic matrix or a transition matrix. A Markov or stochastic matrix is called regular if some power of A has all entries strictly positive. A vector $\mathbf{v} \in \mathbb{R}^n$, is a steady state if $A\mathbf{v} = \mathbf{v}$.

Lemma 10.1.2 *The property of being a stochastic matrix is preserved by taking products. It is also true if the sum is of the form $\sum_j a_{ij} = 1$.*

Proof: Suppose the sum over a row equals 1 for A and B. Then letting the entries be denoted by (a_{ij}) and (b_{ij}) respectively and the entries of AB by (c_{ij}),

$$\sum_i c_{ij} = \sum_i \sum_k a_{ik} b_{kj} = \sum_k \sum_i a_{ik} b_{kj} = \sum_k b_{kj} = 1$$

It is obvious that when the product is taken, if each $a_{ij}, b_{ij} \geq 0$, then the same will be true of sums of products of these numbers. Similar reasoning works for the assumption that $\sum_j a_{ij} = 1$. ■

The following theorem is convenient for showing the existence of limits.

Theorem 10.1.3 *Let A be a real $p \times p$ matrix having the properties*

1. $a_{ij} \geq 0$
2. *Either* $\sum_{i=1}^p a_{ij} = 1$ *or* $\sum_{j=1}^p a_{ij} = 1$.
3. *The distinct eigenvalues of A are $\{1, \lambda_2, \ldots, \lambda_m\}$ where each $|\lambda_j| < 1$.*

Then $\lim_{n\to\infty} A^n = A_\infty$ exists in the sense that $\lim_{n\to\infty} a_{ij}^n = a_{ij}^\infty$, the ij^{th} entry A_∞. ***Here a_{ij}^n denotes the ij^{th} entry of*** *A^n. Also, if $\lambda = 1$ has algebraic multiplicity r, then the Jordan block corresponding to $\lambda = 1$ is just the $r \times r$ identity.*

Proof. By the existence of the Jordan form for A, it follows that there exists an invertible matrix P such that

$$P^{-1}AP = \begin{pmatrix} I+N & & & \\ & J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{r_m}(\lambda_m) \end{pmatrix} = J$$

where I is $r \times r$ for r the multiplicity of the eigenvalue 1 and N is a nilpotent matrix for which $N^r = 0$. I will show that because of Condition 2, $N = 0$.

First of all,

$$J_{r_i}(\lambda_i) = \lambda_i I + N_i$$

where N_i satisfies $N_i^{r_i} = 0$ for some $r_i > 0$. It is clear that $N_i(\lambda_i I) = (\lambda_i I) N$ and so

$$(J_{r_i}(\lambda_i))^n = \sum_{k=0}^{n} \binom{n}{k} N^k \lambda_i^{n-k} = \sum_{k=0}^{r} \binom{n}{k} N^k \lambda_i^{n-k}$$

which converges to 0 due to the assumption that $|\lambda_i| < 1$. There are finitely many terms and a typical one is a matrix whose entries are no larger than an expression of the form

$$|\lambda_i|^{n-k} C_k n(n-1)\cdots(n-k+1) \leq C_k |\lambda_i|^{n-k} n^k$$

which converges to 0 because, by the root test, the series $\sum_{n=1}^\infty |\lambda_i|^{n-k} n^k$ converges. Thus for each $i = 2, \ldots, p$,

$$\lim_{n\to\infty} (J_{r_i}(\lambda_i))^n = 0.$$

By Condition 2, if a_{ij}^n denotes the ij^{th} entry of A^n, then either

$$\sum_{i=1}^p a_{ij}^n = 1 \text{ or } \sum_{j=1}^p a_{ij}^n = 1,\ a_{ij}^n \geq 0.$$

This follows from Lemma 10.1.2. It is obvious each $a_{ij}^n \geq 0$, and so the entries of A^n must be bounded independent of n.

It follows easily from

$$\overbrace{P^{-1}APP^{-1}APP^{-1}AP\cdots P^{-1}AP}^{n \text{ times}} = P^{-1}A^nP$$

that

$$P^{-1}A^nP = J^n \tag{10.1}$$

Hence J^n must also have bounded entries as $n \to \infty$. However, this requirement is incompatible with an assumption that $N \neq 0$.

If $N \neq 0$, then $N^s \neq 0$ but $N^{s+1} = 0$ for some $1 \leq s \leq r$. Then

$$(I+N)^n = I + \sum_{k=1}^{s} \binom{n}{k} N^k$$

One of the entries of N^s is nonzero by the definition of s. Let this entry be n_{ij}^s. Then this implies that one of the entries of $(I+N)^n$ is of the form $\binom{n}{s} n_{ij}^s$. This entry dominates the ij^{th} entries of $\binom{n}{k} N^k$ for all $k < s$ because

$$\lim_{n\to\infty} \binom{n}{s} / \binom{n}{k} = \infty$$

Therefore, the entries of $(I+N)^n$ cannot all be bounded. From block multiplication,

$$P^{-1}A^nP = \begin{pmatrix} (I+N)^n & & & \\ & (J_{r_2}(\lambda_2))^n & & \\ & & \ddots & \\ & & & (J_{r_m}(\lambda_m))^n \end{pmatrix}$$

and this is a contradiction because entries are bounded on the left and unbounded on the right.

Since $N = 0$, the above equation implies $\lim_{n\to\infty} A^n$ exists and equals

$$P \begin{pmatrix} I & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} P^{-1} \blacksquare$$

Are there examples which will cause the eigenvalue condition of this theorem to hold? The following lemma gives such a condition. It turns out that if $a_{ij} > 0$, not just ≥ 0, then the eigenvalue condition of the above theorem is valid.

Lemma 10.1.4 *Suppose $A = (a_{ij})$ is a stochastic matrix. Then $\lambda = 1$ is an eigenvalue. If $a_{ij} > 0$ for all i, j, then if μ is an eigenvalue of A, either $|\mu| < 1$ or $\mu = 1$.*

Proof: First consider the claim that 1 is an eigenvalue. By definition,

$$\sum_i 1 a_{ij} = 1$$

and so $A^T\mathbf{v} = \mathbf{v}$ where $\mathbf{v} = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}^T$. Since A, A^T have the same eigenvalues, this shows 1 is an eigenvalue. Suppose then that μ is an eigenvalue. Is $|\mu| < 1$ or $\mu = 1$? Let $\mathbf{v}$ be an eigenvector for A^T and let $|v_i|$ be the largest of the $|v_j|$.

$$\mu v_i = \sum_j a_{ji} v_j$$

and now multiply both sides by $\overline{\mu v_i}$ to obtain

$$\begin{aligned} |\mu|^2 |v_i|^2 &= \sum_j a_{ji} v_j \overline{\mu v_i} = \sum_j a_{ji} \operatorname{Re}\left(v_j \overline{\mu v_i}\right) \\ &\leq \sum_j a_{ji} |v_i|^2 |\mu| = |\mu| |v_i|^2 \end{aligned}$$

Therefore, $|\mu| \leq 1$. If $|\mu| = 1$, then equality must hold in the above, and so $v_j \overline{v_i \mu}$ must be real and nonnegative for each j. In particular, this holds for $j = i$ which shows $\overline{\mu}$ is real and nonnegative. Thus, in this case, $\mu = 1$ because $\bar{\mu} = \mu$ is nonnegative and equal to 1. The only other case is where $|\mu| < 1$. ■

Lemma 10.1.5 *Let A be any Markov matrix and let $\mathbf{v}$ be a vector having all its components non negative with $\sum_i v_i = c$. Then if $\mathbf{w} = A\mathbf{v}$, it follows that $w_i \geq 0$ for all i and $\sum_i w_i = c$.*

Proof: From the definition of $\mathbf{w}$,

$$w_i \equiv \sum_j a_{ij} v_j \geq 0.$$

Also

$$\sum_i w_i = \sum_i \sum_j a_{ij} v_j = \sum_j \sum_i a_{ij} v_j = \sum_j v_j = c. \blacksquare$$

The following theorem about limits is now easy to obtain.

Theorem 10.1.6 *Suppose A is a Markov matrix in which $a_{ij} > 0$ for all i, j and suppose $\mathbf{w}$ is a vector. Then for each i,*

$$\lim_{k \to \infty} \left(A^k \mathbf{w}\right)_i = v_i$$

where $A\mathbf{v} = \mathbf{v}$. In words, $A^k \mathbf{w}$ always converges to a steady state. In addition to this, if the vector $\mathbf{w}$ satisfies $w_i \geq 0$ for all i and $\sum_i w_i = c$, then the vector $\mathbf{v}$ will also satisfy the conditions, $v_i \geq 0$, $\sum_i v_i = c$.

Proof: By Lemma 10.1.4, since each $a_{ij} > 0$, the eigenvalues are either 1 or have absolute value less than 1. Therefore, the claimed limit exists by Theorem 10.1.3. The assertion that the components are nonnegative and sum to c follows from Lemma 10.1.5. That $A\mathbf{v} = \mathbf{v}$ follows from

$$\mathbf{v} = \lim_{n\to\infty} A^n \mathbf{w} = \lim_{n\to\infty} A^{n+1}\mathbf{w} = A \lim_{n\to\infty} A^n \mathbf{w} = A\mathbf{v}. \blacksquare$$

It is not hard to generalize the conclusion of this theorem to regular Markov processes.

Corollary 10.1.7 *Suppose A is a regular Markov matrix, one for which the entries of A^k are all positive for some k, and suppose $\mathbf{w}$ is a vector. Then for each i,*

$$\lim_{n\to\infty} (A^n\mathbf{w})_i = v_i$$

where $A\mathbf{v} = \mathbf{v}$. In words, $A^n\mathbf{w}$ always converges to a steady state. In addition to this, if the vector $\mathbf{w}$ satisfies $w_i \geq 0$ for all i and $\sum_i w_i = c$, Then the vector $\mathbf{v}$ will also satisfy the conditions $v_i \geq 0$, $\sum_i v_i = c$.

Proof: Let the entries of A^k be all positive for some k. Now suppose that $a_{ij} \geq 0$ for all i, j and $A = (a_{ij})$ is a Markov matrix. Then if $B = (b_{ij})$ is a Markov matrix with $b_{ij} > 0$ for all ij, it follows that BA is a Markov matrix which has strictly positive entries. This is because the ij^{th} entry of BA is

$$\sum_k b_{ik} a_{kj} > 0,$$

Thus, from Lemma 10.1.4, A^k has an eigenvalue equal to 1 for all k sufficiently large, and all the other eigenvalues have absolute value strictly less than 1. The same must be true of A. If $\mathbf{v} \neq \mathbf{0}$ and $A\mathbf{v} = \lambda\mathbf{v}$ and $|\lambda| = 1$, then $A^k\mathbf{v} = \lambda^k\mathbf{v}$ and so, by Lemma 10.1.4, $\lambda^m = 1$ if $m \geq k$. Thus

$$1 = \lambda^{k+1} = \lambda^k \lambda = \lambda$$

By Theorem 10.1.3, $\lim_{n\to\infty} A^n\mathbf{w}$ exists. The rest follows as in Theorem 10.1.6. $\blacksquare$

10.2 Migration Matrices

Definition 10.2.1 *Let n locations be denoted by the numbers $1, 2, \cdots, n$. Also suppose it is the case that each year a_{ij} denotes the proportion of residents in location j which move to location i. Also suppose no one escapes or emigrates from without these n locations. This last assumption requires $\sum_i a_{ij} = 1$. Thus (a_{ij}) is a Markov matrix referred to as a migration matrix.*

If $\mathbf{v} = (x_1, \cdots, x_n)^T$ where x_i is the population of location i at a given instant, you obtain the population of location i one year later by computing $\sum_j a_{ij}x_j = (A\mathbf{v})_i$. Therefore, the population of location i after k years is $\left(A^k\mathbf{v}\right)_i$. Furthermore, Corollary 10.1.7 can be used to predict in the case where A is regular what the long time population will be for the given locations.

As an example of the above, consider the case where $n = 3$ and the migration matrix is of the form

$$\begin{pmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{pmatrix}.$$

Now

$$\begin{pmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{pmatrix}^2 = \begin{pmatrix} .38 & .02 & .15 \\ .28 & .64 & .02 \\ .34 & .34 & .83 \end{pmatrix}$$

and so the Markov matrix is regular. Therefore, $\left(A^k\mathbf{v}\right)_i$ will converge to the i^{th} component of a steady state. It follows the steady state can be obtained from solving the system

$$
\begin{aligned}
.6x + .1z &= x \\
.2x + .8y &= y \\
.2x + .2y + .9z &= z
\end{aligned}
$$

along with the stipulation that the sum of x, y, and z must equal the constant value present at the beginning of the process. The solution to this system is

$$\{y = x, z = 4x, x = x\}.$$

If the total population at the beginning is 150,000, then you solve the following system

$$y = x,\ z = 4x,\ x + y + z = 150000$$

whose solution is easily seen to be $\{x = 25\,000, z = 100\,000, y = 25\,000\}$. Thus, after a long time there would be about four times as many people in the third location as in either of the other two.

10.3 Absorbing States

There is a different kind of Markov process containing so called absorbing states which result in transition matrices which are not regular. However, Theorem 10.1.3 may still apply. One such example is the Gambler's ruin problem. There is a total amount of money denoted by b. The Gambler starts with an amount $j > 0$ and gambles till he either loses everything or gains everything. He does this by playing a game in which he wins with probability p and loses with probability q. When he wins, the amount of money he has increases by 1 and when he loses, the amount of money he has decreases by 1. Thus the states are the integers from 0 to b. Let p_{ij} denote the probability that the gambler has i at the end of a game given that he had j at the beginning. Let p_{ij}^n denote the probability that the gambler has i after n games given that he had j initially. Thus

$$p_{ij}^{n+1} = \sum_k p_{ik} p_{kj}^n,$$

and so p_{ij}^n is the ij^{th} entry of P^n where P is the transition matrix. The above description indicates that this transition probability matrix is of the form

$$P = \begin{pmatrix} 1 & q & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & 0 \\ 0 & p & \ddots & q & \vdots \\ \vdots & & \ddots & 0 & 0 \\ 0 & \cdots & 0 & p & 1 \end{pmatrix} \tag{10.2}$$

The absorbing states are 0 and b. In the first, the gambler has lost everything and hence has nothing else to gamble, so the process stops. In the second, he has won everything and there is nothing else to gain, so again the process stops.

Consider the eigenvalues of this matrix.

Lemma 10.3.1 *Let $p, q > 0$ and $p + q = 1$. Then the eigenvalues of*

$$\begin{pmatrix} 0 & q & 0 & \cdots & 0 \\ p & 0 & q & \cdots & 0 \\ 0 & p & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & q \\ 0 & \vdots & 0 & p & 0 \end{pmatrix}$$

have absolute value less than 1.

Proof: By Gerschgorin's theorem, (See Page 175) if λ is an eigenvalue, then $|\lambda| \leq 1$. Now suppose $\mathbf{v}$ is an eigenvector for λ. Then

$$A\mathbf{v} = \begin{pmatrix} qv_2 \\ pv_1 + qv_3 \\ \vdots \\ pv_{n-2} + qv_n \\ pv_{n-1} \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}.$$

Suppose $|\lambda| = 1$. Let v_k be the first nonzero entry. Then

$$qv_{k+1} = \lambda v_k$$

and so $|v_{k+1}| > |v_k|$. If $\{|v_j|\}_{j=k}^{m}$ is increasing for some $m > k$, then

$$p\,|v_{m-1}| + q\,|v_m| \geq |pv_{m-2} + qv_m| = |\lambda v_{m-1}| = |v_{m-1}|$$

and so $q\,|v_m| \geq q\,|v_{m-1}|$. Thus by induction, the sequence is increasing. Hence $|v_n| \geq |v_{n-1}| > 0$. However, the last line states that $p\,|v_{n-1}| = |v_n|$ which requires that $|v_{n-1}| > |v_n|$, a contradiction. ■

Now consider the eigenvalues of 10.2. For P given there,

$$P - \lambda I = \begin{pmatrix} 1-\lambda & q & 0 & \cdots & 0 \\ 0 & -\lambda & \ddots & & 0 \\ 0 & p & \ddots & q & \vdots \\ \vdots & & \ddots & -\lambda & 0 \\ 0 & \cdots & 0 & p & 1-\lambda \end{pmatrix}$$

and so, expanding the determinant of the matrix along the first column and then along the last column yields

$$(1-\lambda)^2 \det \begin{pmatrix} -\lambda & q & & \\ p & \ddots & \ddots & \\ & \ddots & -\lambda & q \\ & & p & -\lambda \end{pmatrix}.$$

The roots of the polynomial after $(1-\lambda)^2$ have absolute value less than 1 because they are just the eigenvalues of a matrix of the sort in Lemma 10.3.1. It follows that the conditions of Theorem 10.1.3 apply and therefore, $\lim_{n\to\infty} P^n$ exists. ■

Of course, the above transition matrix, models many other kinds of problems. It is called a Markov process with two absorbing states, sometimes a random walk with two absorbing states.

It is interesting to find the probability that the gambler loses all his money. This is given by $\lim_{n\to\infty} p_{0j}^n$.From the transition matrix for the gambler's ruin problem, it follows that

$$\begin{aligned} p_{0j}^n &= \sum_k p_{0k}^{n-1} p_{kj} = q p_{0(j-1)}^{n-1} + p p_{0(j+1)}^{n-1} \text{for } j \in [1, b-1], \\ p_{00}^n &= 1, \text{ and } p_{0b}^n = 0. \end{aligned}$$

Assume here that $p \neq q$. Now it was shown above that $\lim_{n\to\infty} p_{0j}^n$ exists. Denote by P_j this limit. Then the above becomes much simpler if written as

$$P_j = q P_{j-1} + p P_{j+1} \text{ for } j \in [1, b-1], \tag{10.3}$$

$$P_0 = 1 \text{ and } P_b = 0. \tag{10.4}$$

It is only required to find a solution to the above difference equation with boundary conditions. To do this, look for a solution in the form $P_j = r^j$and use the difference equation with boundary conditions to find the correct values of r. Thus you need

$$r^j = q r^{j-1} + p r^{j+1}$$

and so to find r you need to have $pr^2 - r + q = 0$, and so the solutions for r are $r =$

$$\frac{1}{2p}\left(1+\sqrt{1-4pq}\right), \ \frac{1}{2p}\left(1-\sqrt{1-4pq}\right)$$

Now

$$\sqrt{1-4pq} = \sqrt{1-4p(1-p)} = \sqrt{1-4p+4p^2} = 1-2p.$$

Thus the two values of r simplify to

$$\frac{1}{2p}(1+1-2p) = \frac{q}{p}, \quad \frac{1}{2p}(1-(1-2p)) = 1$$

Therefore, for any choice of $C_i, i = 1, 2$,

$$C_1 + C_2 \left(\frac{q}{p}\right)^j$$

will solve the difference equation. Now choose C_1, C_2 to satisfy the boundary conditions 10.4. Thus you need to have

$$C_1 + C_2 = 1, \ C_1 + C_2 \left(\frac{q}{p}\right)^b = 0$$

It follows that

$$C_2 = \frac{p^b}{p^b - q^b}, \quad C_1 = \frac{q^b}{q^b - p^b}$$

Thus $P_j =$

$$\frac{q^b}{q^b - p^b} + \frac{p^b}{p^b - q^b}\left(\frac{q}{p}\right)^j = \frac{q^b}{q^b - p^b} - \frac{p^{b-j}q^j}{q^b - p^b} = \frac{q^j\left(q^{b-j} - p^{b-j}\right)}{q^b - p^b}$$

To find the solution in the case of a fair game, one could take the $\lim_{p\to 1/2}$ of the above solution. Taking this limit, you get

$$P_j = \frac{b-j}{b}.$$

You could also verify directly in the case where $p = q = 1/2$ in 10.3 and 10.4 that $P_j = 1$ and $P_j = j$ are two solutions to the difference equation and proceeding as before.

10.4 Exercises

1. Suppose the migration matrix for three locations is

$$\begin{pmatrix} .5 & 0 & .3 \\ .3 & .8 & 0 \\ .2 & .2 & .7 \end{pmatrix}.$$

Find a comparison for the populations in the three locations after a long time.

2. Show that if $\sum_i a_{ij} = 1$, then if $A = (a_{ij})$, then the sum of the entries of $A\mathbf{v}$ equals the sum of the entries of $\mathbf{v}$. Thus it does not matter whether $a_{ij} \geq 0$ for this to be so.

3. If A satisfies the conditions of the above problem, can it be concluded that $\lim_{n\to\infty} A^n$ exists?

4. Give an example of a non regular Markov matrix which has an eigenvalue equal to -1.

5. Show that when a Markov matrix is non defective, all of the above theory can be proved very easily. In particular, prove the theorem about the existence of $\lim_{n\to\infty} A^n$ if the eigenvalues are either 1 or have absolute value less than 1.

6. Find a formula for A^n where

$$A = \begin{pmatrix} \frac{5}{2} & -\frac{1}{2} & 0 & -1 \\ 5 & 0 & 0 & -4 \\ \frac{7}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{5}{2} \\ \frac{7}{2} & -\frac{1}{2} & 0 & -2 \end{pmatrix}$$

Does $\lim_{n\to\infty} A^n$ exist? Note that all the rows sum to 1. **Hint:** This matrix is similar to a diagonal matrix. The eigenvalues are $1, -1, \frac{1}{2}, \frac{1}{2}$.

7. Find a formula for A^n where

$$A = \begin{pmatrix} 2 & -\frac{1}{2} & \frac{1}{2} & -1 \\ 4 & 0 & 1 & -4 \\ \frac{5}{2} & -\frac{1}{2} & 1 & -2 \\ 3 & -\frac{1}{2} & \frac{1}{2} & -2 \end{pmatrix}$$

Note that the rows sum to 1 in this matrix also. **Hint:** This matrix is not similar to a diagonal matrix but you can find the Jordan form and consider this in order to obtain a formula for this product. The eigenvalues are $1, -1, \frac{1}{2}, \frac{1}{2}$.

8. Find $\lim_{n\to\infty} A^n$ if it exists for the matrix

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{2} & 0 \\ \frac{3}{2} & \frac{3}{2} & \frac{3}{2} & 1 \end{pmatrix}$$

The eigenvalues are $\frac{1}{2}, 1, 1, 1$.

9. Give an example of a matrix A which has eigenvalues which are either equal to 1,−1, or have absolute value strictly less than 1 but which has the property that $\lim_{n\to\infty} A^n$ does not exist.

10. If A is an $n \times n$ matrix such that all the eigenvalues have absolute value less than 1, show $\lim_{n\to\infty} A^n = 0$.

11. Find an example of a 3×3 matrix A such that $\lim_{n\to\infty} A^n$ does not exist but $\lim_{r\to\infty} A^{5r}$ does exist.

12. If A is a Markov matrix and B is similar to A, does it follow that B is also a Markov matrix?

13. In Theorem 10.1.3 suppose everything is unchanged except that you assume either $\sum_j a_{ij} \leq 1$ or $\sum_i a_{ij} \leq 1$. Would the same conclusion be valid? What if you don't insist that each $a_{ij} \geq 0$? Would the conclusion hold in this case?

14. Let V be an n dimensional vector space and let $\mathbf{x} \in V$ and $\mathbf{x} \neq \mathbf{0}$. Consider $\beta_{\mathbf{x}} \equiv \mathbf{x}, A\mathbf{x}, \cdots, A^{m-1}\mathbf{x}$ where

$$A^m\mathbf{x} \in \text{span}\left(\mathbf{x}, A\mathbf{x}, \cdots, A^{m-1}\mathbf{x}\right)$$

and m is the smallest such that the above inclusion in the span takes place. Show that $\left\{\mathbf{x},A\mathbf{x},\cdots,A^{m-1}\mathbf{x}\right\}$ must be linearly independent. Next suppose $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ is a basis for V. Consider $\beta_{\mathbf{v}_i}$ as just discussed, having length m_i. Thus $A^{m_i}\mathbf{v}_i$ is a linearly combination of $\mathbf{v}_i,A\mathbf{v}_i,\cdots,A^{m-1}\mathbf{v}_i$ for m as small as possible. Let $p_{\mathbf{v}_i}(\lambda)$ be the monic polynomial which expresses this linear combination. Thus $p_{\mathbf{v}_i}(A)\mathbf{v}_i = 0$ and the degree of $p_{\mathbf{v}_i}(\lambda)$ is as small as possible for this to take place. Show that the minimal polynomial for A must be the monic polynomial which is the least common multiple of these polynomials $p_{\mathbf{v}_i}(\lambda)$.

15. If A is a complex Hermitian $n\times n$ matrix which has all eigenvalues nonnegative, show that there exists a complex Hermitian matrix B such that $BB = A$.

16. ↑Suppose A,B are $n\times n$ real Hermitian matrices and they both have all nonnegative eigenvalues. Show that $\det(A+B)\geq\det(A)+\det(B)$. **Hint:** Use the above problem and the Cauchy Binet theorem. Let $P^2 = A, Q^2 = B$ where P,Q are Hermitian and nonnegative. Then
$$A+B=\begin{pmatrix} P & Q\end{pmatrix}\begin{pmatrix} P \\ Q\end{pmatrix}.$$

17. Suppose $B=\begin{pmatrix}\alpha & \mathbf{c}^* \\ \mathbf{b} & A\end{pmatrix}$ is an $(n+1)\times(n+1)$ Hermitian nonnegative matrix where α is a scalar and A is $n\times n$. Show that α must be real, $\mathbf{c}=\mathbf{b}$, and $A = A^*, A$ is nonnegative, and that if $\alpha = 0$, then $\mathbf{b} = \mathbf{0}$. Otherwise, $\alpha > 0$.

18. ↑If A is an $n\times n$ complex Hermitian and nonnegative matrix, show that there exists an upper triangular matrix B such that $B^*B = A$. **Hint:** Prove this by induction. It is obviously true if $n = 1$. Now if you have an $(n+1)\times(n+1)$ Hermitian nonnegative matrix, then from the above problem, it is of the form $\begin{pmatrix}\alpha^2 & \alpha\mathbf{b}^* \\ \alpha\mathbf{b} & A\end{pmatrix}$, α real.

19. ↑ Suppose A is a nonnegative Hermitian matrix (all eigenvalues are nonnegative) which is partitioned as
$$A=\begin{pmatrix}A_{11} & A_{12} \\ A_{21} & A_{22}\end{pmatrix}$$
where A_{11},A_{22} are square matrices. Show that $\det(A)\leq\det(A_{11})\det(A_{22})$. **Hint:** Use the above problem to factor A getting
$$A=\begin{pmatrix}B_{11}^* & 0^* \\ B_{12}^* & B_{22}^*\end{pmatrix}\begin{pmatrix}B_{11} & B_{12} \\ 0 & B_{22}\end{pmatrix}$$
Next argue that $A_{11} = B_{11}^*B_{11}, A_{22} = B_{12}^*B_{12} + B_{22}^*B_{22}$. Use the Cauchy Binet theorem to argue that $\det(A_{22}) = \det(B_{12}^*B_{12} + B_{22}^*B_{22}) \geq \det(B_{22}^*B_{22})$. Then explain why
$$\begin{aligned}\det(A) &= \det(B_{11}^*)\det(B_{22}^*)\det(B_{11})\det(B_{22}) \\ &= \det(B_{11}^*B_{11})\det(B_{22}^*B_{22})\end{aligned}$$

20. ↑ Prove the inequality of Hadamard. If A is a Hermitian matrix which is nonnegative (all eigenvalues are nonnegative), then $\det(A)\leq\prod_i A_{ii}$.

Chapter 11

Inner Product Spaces

11.1 General Theory

It is assumed here that the field of scalars is either $\mathbb{R}$ or $\mathbb{C}$. The usual example of an inner product space is $\mathbb{C}^n$ or $\mathbb{R}^n$ as described earlier. However, there are many other inner product spaces and the topic is of such importance that it seems appropriate to discuss the general theory of these spaces.

Definition 11.1.1 *A vector space X is said to be a normed linear space if there exists a function, denoted by $|\cdot| : X \to [0, \infty)$ which satisfies the following axioms.*

1. $|x| \geq 0$ *for all* $x \in X$, *and* $|x| = 0$ *if and only if* $x = 0$.
2. $|ax| = |a| \, |x|$ *for all* $a \in \mathbb{F}$.
3. $|x + y| \leq |x| + |y|$.

This function $|\cdot|$ is called a norm.

The notation $||x||$ is also often used. Not all norms are created equal. There are many geometric properties which they may or may not possess. There is also a concept called an inner product which is discussed next. It turns out that the best norms come from an inner product.

Definition 11.1.2 *A mapping $(\cdot, \cdot) : V \times V \to \mathbb{F}$ is called an inner product if it satisfies the following axioms.*

1. $(x, y) = \overline{(y, x)}$.
2. $(x, x) \geq 0$ *for all* $x \in V$ *and equals zero if and only if* $x = 0$.
3. $(ax + by, z) = a(x, z) + b(y, z)$ *whenever* $a, b \in \mathbb{F}$.

Note that 2 and 3 imply $(x, ay + bz) = \overline{a}(x, y) + \overline{b}(x, z)$.
Then a norm is given by

$$(x, x)^{1/2} \equiv |x| .$$

It remains to verify this really is a norm.

Definition 11.1.3 *A normed linear space in which the norm comes from an inner product as just described is called an inner product space.*

Example 11.1.4 *Let $V = \mathbb{C}^n$ with the inner product given by*

$$(\mathbf{x}, \mathbf{y}) \equiv \sum_{k=1}^{n} x_k \overline{y}_k.$$

This is an example of a complex inner product space already discussed.

Example 11.1.5 *Let $V = \mathbb{R}^n$,*

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \equiv \sum_{j=1}^{n} x_j y_j.$$

This is an example of a real inner product space.

Example 11.1.6 *Let V be any finite dimensional vector space and let $\{v_1, \cdots, v_n\}$ be a basis. Decree that*

$$(v_i, v_j) \equiv \delta_{ij} \equiv \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

and define the inner product by

$$(x, y) \equiv \sum_{i=1}^{n} x^i \overline{y^i}$$

where

$$x = \sum_{i=1}^{n} x^i v_i, \; y = \sum_{i=1}^{n} y^i v_i.$$

The above is well defined because $\{v_1, \cdots, v_n\}$ is a basis. Thus the components x_i associated with any given $x \in V$ are uniquely determined.

This example shows there is no loss of generality when studying finite dimensional vector spaces with field of scalars $\mathbb{R}$ or $\mathbb{C}$ in assuming the vector space is actually an inner product space. The following theorem was presented earlier with slightly different notation.

Theorem 11.1.7 *(Cauchy Schwarz) In any inner product space*

$$|(x, y)| \leq |x||y|.$$

where $|x| \equiv (x, x)^{1/2}$.

Proof: Let $\omega \in \mathbb{C}$, $|\omega| = 1$, and $\overline{\omega}(x, y) = |(x, y)| = \text{Re}(x, y\omega)$. Let

$$F(t) = (x + ty\omega, x + t\omega y).$$

Then from the axioms of the inner product,

$$F(t) = |x|^2 + 2t\,\mathrm{Re}(x, \omega y) + t^2|y|^2 \geq 0.$$

This yields

$$|x|^2 + 2t|(x, y)| + t^2|y|^2 \geq 0.$$

If $|y| = 0$, then the inequality requires that $|(x, y)| = 0$ since otherwise, you could pick large negative t and contradict the inequality. If $|y| > 0$, it follows from the quadratic formula that

$$4|(x, y)|^2 - 4|x|^2|y|^2 \leq 0. \blacksquare$$

Earlier it was claimed that the inner product defines a norm. In this next proposition this claim is proved.

Proposition 11.1.8 *For an inner product space,* $|x| \equiv (x, x)^{1/2}$ *does specify a norm.*

Proof: All the axioms are obvious except the triangle inequality. To verify this,

$$\begin{aligned} |x + y|^2 &\equiv (x + y, x + y) \equiv |x|^2 + |y|^2 + 2\,\mathrm{Re}\,(x, y) \\ &\leq |x|^2 + |y|^2 + 2\,|(x, y)| \\ &\leq |x|^2 + |y|^2 + 2\,|x|\,|y| = (|x| + |y|)^2. \blacksquare \end{aligned}$$

The best norms of all are those which come from an inner product because of the following identity which is known as the parallelogram identity.

Proposition 11.1.9 *If* $(V, (\cdot, \cdot))$ *is an inner product space then for* $|x| \equiv (x, x)^{1/2}$*, the following identity holds.*

$$|x + y|^2 + |x - y|^2 = 2\,|x|^2 + 2\,|y|^2 .$$

It turns out that the validity of this identity is equivalent to the existence of an inner product which determines the norm as described above. These sorts of considerations are topics for more advanced courses on functional analysis.

Definition 11.1.10 *A basis for an inner product space,* $\{u_1, \cdots, u_n\}$ *is an orthonormal basis if*

$$(u_k, u_j) = \delta_{kj} \equiv \begin{cases} 1 \text{ if } k = j \\ 0 \text{ if } k \neq j \end{cases}.$$

Note that if a list of vectors satisfies the above condition for being an orthonormal set, then the list of vectors is automatically linearly independent. To see this, suppose

$$\sum_{j=1}^{n} c^j u_j = 0$$

Then taking the inner product of both sides with u_k,

$$0 = \sum_{j=1}^{n} c^j\,(u_j, u_k) = \sum_{j=1}^{n} c^j \delta_{jk} = c^k.$$

11.2 The Gram Schmidt Process

Lemma 11.2.1 *Let* X *be an inner product space and let* $\{x_1, \cdots, x_n\}$ *be linearly independent. Then there exists an orthonormal basis for* X*,* $\{u_1, \cdots, u_n\}$ *which has the property that for each* $k \leq n$*, span*$(x_1, \cdots, x_k) = \mathrm{span}\,(u_1, \cdots, u_k)$.

Proof: Let $u_1 \equiv x_1/\,|x_1|$. Thus for $k = 1$, $\mathrm{span}\,(u_1) = \mathrm{span}\,(x_1)$ and $\{u_1\}$ is an orthonormal set. Now suppose for some $k < n$, $u_1, \cdots, u_k$ have been chosen such that $(u_j, u_l) = \delta_{jl}$ and $\mathrm{span}\,(x_1, \cdots, x_k) = \mathrm{span}\,(u_1, \cdots, u_k)$. Then define

$$u_{k+1} \equiv \frac{x_{k+1} - \sum_{j=1}^{k} (x_{k+1}, u_j) u_j}{\left| x_{k+1} - \sum_{j=1}^{k} (x_{k+1}, u_j) u_j \right|}, \tag{11.1}$$

where the denominator is not equal to zero because the x_j form a basis and so

$$x_{k+1} \notin \text{span}(x_1, \cdots, x_k) = \text{span}(u_1, \cdots, u_k)$$

Thus by induction,

$$u_{k+1} \in \text{span}(u_1, \cdots, u_k, x_{k+1}) = \text{span}(x_1, \cdots, x_k, x_{k+1}).$$

Also, $x_{k+1} \in \text{span}(u_1, \cdots, u_k, u_{k+1})$ which is seen easily by solving 11.1 for x_{k+1} and it follows

$$\text{span}(x_1, \cdots, x_k, x_{k+1}) = \text{span}(u_1, \cdots, u_k, u_{k+1}).$$

If $l \leq k$,

$$\begin{aligned}(u_{k+1}, u_l) &= C\left((x_{k+1}, u_l) - \sum_{j=1}^{k} (x_{k+1}, u_j)(u_j, u_l)\right) \\ &= C\left((x_{k+1}, u_l) - \sum_{j=1}^{k} (x_{k+1}, u_j)\delta_{lj}\right) \\ &= C((x_{k+1}, u_l) - (x_{k+1}, u_l)) = 0.\end{aligned}$$

The vectors, $\{u_j\}_{j=1}^n$, generated in this way are therefore an orthonormal basis because each vector has unit length. ■

The process by which these vectors were generated is called the Gram Schmidt process. The following corollary is obtained from the above process.

Corollary 11.2.2 *Let X be a finite dimensional inner product space of dimension n whose basis is $\{u_1, \cdots, u_k, x_{k+1}, \cdots, x_n\}$. Then if $\{u_1, \cdots, u_k\}$ is orthonormal, then the Gram Schmidt process applied to the given list of vectors in order leaves $\{u_1, \cdots, u_k\}$ unchanged.*

Lemma 11.2.3 *Suppose $\{u_j\}_{j=1}^n$ is an orthonormal basis for an inner product space X. Then for all $x \in X$,*

$$x = \sum_{j=1}^{n} (x, u_j) u_j.$$

Proof: Since $\{u_j\}_{j=1}^n$ is a basis, there exist unique scalars $\{\alpha_i\}$ such that

$$x = \sum_{j=1}^{n} \alpha_j u_j$$

It only remains to identify α_k. From the properties of the inner product,

$$(x, u_k) = \sum_{j=1}^{n} \alpha_j (u_j, u_k) = \sum_{j=1}^{n} \alpha_j \delta_{jk} = \alpha_k \; ■$$

The following theorem is of fundamental importance. First note that a subspace of an inner product space is also an inner product space because you can use the same inner product.

Theorem 11.2.4 *Let M be a finite dimensional subspace of X, an inner product space and let $\{e_i\}_{i=1}^m$ be an orthonormal basis for M. Then if $y \in X$ and $w \in M$,*

$$|y - w|^2 = \inf\left\{|y - z|^2 : z \in M\right\} \tag{11.2}$$

if and only if

$$(y - w, z) = 0 \tag{11.3}$$

for all $z \in M$*. Furthermore,*

$$w = \sum_{i=1}^{m} (y, x_i) x_i \tag{11.4}$$

is the unique element of M *which has this property. It is called the orthogonal projection.*

Proof: First we show that if 11.3, then 11.2. Let $z \in M$ be arbitrary. Then

$$|y - z|^2 = |y - w + (w - z)|^2$$
$$= (y - w + (w - z), y - w + (w - z))$$
$$= |y - w|^2 + |z - w|^2 + 2\,\mathrm{Re}\,(y - w, w - z)$$

The last term is given to be 0 and so

$$|y - z|^2 = |y - w|^2 + |z - w|^2$$

which verifies 11.2.

Next suppose 11.2. Is it true that 11.3 follows? Let $z \in M$ be arbitrary and let $|\theta| = 1, \bar{\theta}(x - w, w - z) = |(x - w, w - z)|$. Then let

$$\begin{aligned} p(t) &\equiv |x - w + t\theta(w - z)|^2 = |x - w|^2 + 2\,\mathrm{Re}\,(x - w, t\theta(w - z)) + t^2 |w - z|^2 \\ &= |x - w|^2 + 2\,\mathrm{Re}\, t\bar{\theta}(x - w, (w - z)) + t^2 |w - z|^2 \\ &= |x - w|^2 + 2t |(x - w, (w - z))| + t^2 |w - z|^2 \end{aligned}$$

Then p has a minimum when $t = 0$ and so $p'(0) = 2|(x - w, (w - z))| = 0$ which shows 11.3. This proves the first part of the theorem since z is arbitrary.

It only remains to verify that w given in 11.4 satisfies 11.3 and is the only point of M which does so.

First, could there be two minimizers? Say w_1, w_2 both work. Then by the above characterization of minimizers,

$$\begin{aligned}(x - w_1, w_1 - w_2) &= 0 \\ (x - w_2, w_1 - w_2) &= 0\end{aligned}$$

Subtracting gives $(w_1 - w_2, w_1 - w_2) = 0$. Hence the minimizer is unique.

Finally, it remains to show that the given formula works. Letting $\{e_1, \cdots, e_m\}$ be an orthonormal basis for M, such a thing existing by the Gramm Schmidt process,

$$\begin{aligned}\left(x - \sum_{i=1}^{m} (x, e_i) e_i, e_k\right) &= (x, e_k) - \sum_{i=1}^{m} (x, e_i)(e_i, e_k) \\ &= (x, e_k) - \sum_{i=1}^{m} (x, e_i)\delta_{ik} \\ &= (x, e_k) - (x, e_k) = 0\end{aligned}$$

Since this inner product equals 0 for arbitrary e_k, it follows that

$$\left(x - \sum_{i=1}^{m} (x, e_i) e_i, z\right) = 0$$

for every $z \in M$ because each such z is a linear combination of the e_i. Hence $\sum_{i=1}^{m} (x, e_i) e_i$ is the unique minimizer. ■

Example 11.2.5 *Consider X equal to the continuous functions defined on $[-\pi, \pi]$ and let the inner product be given by*

$$\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$$

It is left to the reader to verify that this is an inner product. Letting e_k be the function $x \to \frac{1}{\sqrt{2\pi}} e^{ikx}$, define

$$M \equiv \text{span}\left(\{e_k\}_{k=-n}^{n}\right).$$

Then you can verify that

$$(e_k, e_m) = \int_{-\pi}^{\pi} \left(\frac{1}{\sqrt{2\pi}} e^{-ikx}\right)\left(\frac{1}{\sqrt{2\pi}} \overline{e^{mix}}\right) dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-k)x} = \delta_{km}$$

then for a given function $f \in X$, the function from M which is closest to f in this inner product norm is

$$g = \sum_{k=-n}^{n} (f, e_k) e_k$$

In this case $(f, e_k) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{ikx} dx$. These are the Fourier coefficients. The above is the n^{th} partial sum of the Fourier series.

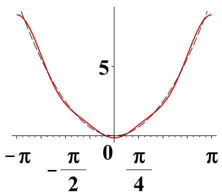
To show how this kind of thing approximates a given function, let $f(x) = x^2$. Let $M = \text{span}\left(\left\{\frac{1}{\sqrt{2\pi}} e^{-ikx}\right\}_{k=-3}^{3}\right)$. Then, doing the computations, you find the closest point is of the form

$$\frac{1}{3}\sqrt{2}\pi^{\frac{5}{2}}\left(\frac{1}{\sqrt{2\pi}}\right) + \sum_{k=1}^{3}\left(\frac{(-1)^k 2}{k^2}\right)\sqrt{2}\sqrt{\pi}\frac{1}{\sqrt{2\pi}} e^{-ikx} + \sum_{k=1}^{3}\left(\frac{(-1)^k 2}{k^2}\right)\sqrt{2}\sqrt{\pi}\frac{1}{\sqrt{2\pi}} e^{ikx}$$

and now simplify to get

$$\frac{1}{3}\pi^2 + \sum_{k=1}^{3} (-1)^k \left(\frac{4}{k^2}\right)\cos kx$$

Then a graph of this along with the graph of $y = x^2$ is given below. In this graph, the dashed graph is of $y = x^2$ and the solid line is the graph of the above Fourier series approximation.



If we had taken the partial sum up to n much bigger, it would have been very hard to distinguish between the graph of the partial sum of the Fourier series and the graph of the function it is approximating. This is in contrast to approximation by Taylor series in which you only get approximation at a point of a function and its derivatives. These are very close near the point of interest but typically fail to approximate the function on the entire interval.

11.3 Riesz Representation Theorem

The next theorem is one of the most important results in the theory of inner product spaces. It is called the Riesz representation theorem.

Theorem 11.3.1 *Let $f \in \mathcal{L}(X, \mathbb{F})$ where X is an inner product space of dimension n. Then there exists a unique $z \in X$ such that for all $x \in X$,*

$$f(x) = (x, z).$$

Proof: First I will verify uniqueness. Suppose z_j works for $j = 1, 2$. Then for all $x \in X$,

$$0 = f(x) - f(x) = (x, z_1 - z_2)$$

and so $z_1 = z_2$.

It remains to verify existence. By Lemma 11.2.1, there exists an orthonormal basis, $\{u_j\}_{j=1}^n$. If there is such a z, then you would need $f(u_j) = (u_j, z)$ and so you would need $\overline{f(u_j)} = (z, u_j)$. Also you must have $z = \sum_i (z, u_j) u_j$. Therefore, define

$$z \equiv \sum_{j=1}^{n} \overline{f(u_j)} u_j.$$

Then using Lemma 11.2.3,

$$\begin{aligned}(x, z) &= \left(x, \sum_{j=1}^{n} \overline{f(u_j)} u_j\right) = \sum_{j=1}^{n} f(u_j)(x, u_j) \\ &= f\left(\sum_{j=1}^{n} (x, u_j) u_j\right) = f(x). \blacksquare\end{aligned}$$

Corollary 11.3.2 *Let $A \in \mathcal{L}(X, Y)$ where X and Y are two inner product spaces of finite dimension. Then there exists a unique $A^* \in \mathcal{L}(Y, X)$ such that*

$$(Ax, y)_Y = (x, A^* y)_X \tag{11.5}$$

for all $x \in X$ and $y \in Y$. The following formula holds

$$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$$

Proof: Let $f_y \in \mathcal{L}(X, \mathbb{F})$ be defined as

$$f_y(x) \equiv (Ax, y)_Y.$$

Then by the Riesz representation theorem, there exists a unique element of X, $A^*(y)$ such that

$$(Ax, y)_Y = (x, A^*(y))_X.$$

It only remains to verify that A^* is linear. Let a and b be scalars. Then for all $x \in X$,

$$(x, A^*(ay_1 + by_2))_X \equiv (Ax, (ay_1 + by_2))_Y$$

$$\equiv \overline{a}\,(Ax, y_1) + \overline{b}\,(Ax, y_2) \equiv$$

$$\overline{a}\,(x, A^*(y_1)) + \overline{b}\,(x, A^*(y_2)) = (x, aA^*(y_1) + bA^*(y_2)).$$

Since this holds for every x, it follows

$$A^*(ay_1 + by_2) = aA^*(y_1) + bA^*(y_2)$$

which shows A^* is linear as claimed.

Consider the last assertion that * is conjugate linear.

$$\begin{aligned} &\left(x, (\alpha A + \beta B)^* y\right) \equiv ((\alpha A + \beta B)\, x, y) \\ = \;& \alpha\,(Ax, y) + \beta\,(Bx, y) = \alpha\,(x, A^* y) + \beta\,(x, B^* y) \\ = \;& (x, \overline{\alpha} A^* y) + \left(x, \overline{\beta} A^* y\right) = \left(x, \left(\overline{\alpha} A^* + \overline{\beta} A^*\right) y\right). \end{aligned}$$

Since x is arbitrary,

$$(\alpha A + \beta B)^* y = \left(\overline{\alpha} A^* + \overline{\beta} A^*\right) y$$

and since this is true for all y,

$$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} A^*. \;\blacksquare$$

Definition 11.3.3 *The linear map, A^* is called the adjoint of A. In the case when $A : X \to X$ and $A = A^*$, A is called a self adjoint map. Such a map is also called Hermitian.*

Theorem 11.3.4 *Let M be an $m \times n$ matrix. Then $M^* = \left(\overline{M}\right)^T$ in words, the transpose of the conjugate of M is equal to the adjoint.*

Proof: Using the definition of the inner product in $\mathbb{C}^n$,

$$(M\mathbf{x},\mathbf{y}) = (\mathbf{x},M^*\mathbf{y}) \equiv \sum_i x_i \overline{\sum_j (M^*)_{ij} y_j} = \sum_{i,j} \overline{(M^*)_{ij}}\, \overline{y_j} x_i.$$

Also

$$(M\mathbf{x},\mathbf{y}) = \sum_j \sum_i M_{ji} \overline{y_j} x_i.$$

Since $\mathbf{x},\mathbf{y}$ are arbitrary vectors, it follows that $M_{ji} = \overline{(M^*)_{ij}}$ and so, taking conjugates of both sides,

$$M^*_{ij} = \overline{M_{ji}} \quad \blacksquare$$

The next theorem is interesting. You have a p dimensional subspace of $\mathbb{F}^n$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Of course this might be "slanted". However, there is a linear transformation Q which preserves distances which maps this subspace to $\mathbb{F}^p$.

Theorem 11.3.5 *Suppose V is a subspace of $\mathbb{F}^n$ having dimension $p \leq n$. Then there exists a $Q \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ such that*

$$QV \subseteq \text{span}(\mathbf{e}_1, \cdots, \mathbf{e}_p)$$

and $|Q\mathbf{x}| = |\mathbf{x}|$ for all $\mathbf{x}$. Also

$$Q^*Q = QQ^* = I.$$

Proof: By Lemma 11.2.1 there exists an orthonormal basis for $V, \{\mathbf{v}_i\}_{i=1}^p$. By using the Gram Schmidt process this may be extended to an orthonormal basis of the whole space $\mathbb{F}^n$,

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_p, \mathbf{v}_{p+1}, \cdots, \mathbf{v}_n\}.$$

Now define $Q \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ by $Q(\mathbf{v}_i) \equiv \mathbf{e}_i$ and extend linearly. If $\sum_{i=1}^n x_i \mathbf{v}_i$ is an arbitrary element of $\mathbb{F}^n$,

$$\left| Q\left(\sum_{i=1}^n x_i \mathbf{v}_i \right) \right|^2 = \left| \sum_{i=1}^n x_i \mathbf{e}_i \right|^2 = \sum_{i=1}^n |x_i|^2 = \left| \sum_{i=1}^n x_i \mathbf{v}_i \right|^2.$$

It remains to verify that $Q^*Q = QQ^* = I$. To do so, let $\mathbf{x},\mathbf{y} \in \mathbb{F}^n$. Then let ω be a complex number such that $|\omega| = 1, \omega(\mathbf{x},Q^*Q\mathbf{y} - \mathbf{y}) = |(\mathbf{x},Q^*Q\mathbf{y} - \mathbf{y})|$.

$$(Q(\omega\mathbf{x} + \mathbf{y}), Q(\omega\mathbf{x} + \mathbf{y})) = (\omega\mathbf{x} + \mathbf{y}, \omega\mathbf{x} + \mathbf{y}).$$

Thus

$$|Q\mathbf{x}|^2 + |Q\mathbf{y}|^2 + 2\,\text{Re}\,\omega(Q\mathbf{x},Q\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\,\text{Re}\,\omega(\mathbf{x},\mathbf{y})$$

and since Q preserves norms, it follows that for all $\mathbf{x},\mathbf{y} \in \mathbb{F}^n$,

$$\text{Re}\,\omega(Q\mathbf{x},Q\mathbf{y}) = \text{Re}\,\omega(\mathbf{x},Q^*Q\mathbf{y}) = \omega\,\text{Re}(\mathbf{x},\mathbf{y}).$$

Thus

$$0 = \text{Re}\,\omega((\mathbf{x},Q^*Q\mathbf{y}) - (\mathbf{x},\mathbf{y})) = \text{Re}\,\omega(\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y}) = |(\mathbf{x},Q^*Q\mathbf{y} - \mathbf{y})|$$

$$\text{Re}(\mathbf{x},Q^*Q\mathbf{y} - \mathbf{y}) = 0 \tag{11.6}$$

for all $\mathbf{x},\mathbf{y}$. Letting $\mathbf{x} = Q^*Q\mathbf{y} - \mathbf{y}$, it follows $Q^*Q\mathbf{y} = \mathbf{y}$. Similarly $QQ^* = I$. $\blacksquare$

Note that is is actually shown that $QV = \text{span}(\mathbf{e}_1, \cdots, \mathbf{e}_p)$ and that in case $p = n$ one obtains that a linear transformation which maps an orthonormal basis to an orthonormal basis is unitary.

11.4 The Tensor Product Of Two Vectors

Definition 11.4.1 *Let X and Y be inner product spaces and let $x \in X$ and $y \in Y$. Define the tensor product of these two vectors, $y \otimes x$, an element of $\mathcal{L}(X,Y)$ by*

$$y \otimes x(u) \equiv y(u,x)_X.$$

This is also called a rank one transformation because the image of this transformation is contained in the span of the vector, y.

The verification that this is a linear map is left to you. Be sure to verify this! The following lemma has some of the most important properties of this linear transformation.

Lemma 11.4.2 *Let X, Y, Z be inner product spaces. Then for α a scalar,*

$$(\alpha (y \otimes x))^* = \overline{\alpha} x \otimes y \tag{11.7}$$

$$(z \otimes y_1)(y_2 \otimes x) = (y_2, y_1) z \otimes x \tag{11.8}$$

Proof: Let $u \in X$ and $v \in Y$. Then

$$(\alpha (y \otimes x) u, v) = (\alpha (u, x) y, v) = \alpha (u, x)(y, v)$$

and

$$(u, \overline{\alpha} x \otimes y (v)) = (u, \overline{\alpha} (v, y) x) = \alpha (y, v)(u, x).$$

Therefore, this verifies 11.7.

To verify 11.8, let $u \in X$.

$$(z \otimes y_1)(y_2 \otimes x)(u) = (u, x)(z \otimes y_1)(y_2) = (u, x)(y_2, y_1) z$$

and

$$(y_2, y_1) z \otimes x (u) = (y_2, y_1)(u, x) z.$$

Since the two linear transformations on both sides of 11.8 give the same answer for every $u \in X$, it follows the two transformations are the same. ■

Definition 11.4.3 *Let X, Y be two vector spaces. Then define for $A, B \in \mathcal{L}(X, Y)$ and $\alpha \in \mathbb{F}$, new elements of $\mathcal{L}(X, Y)$ denoted by $A + B$ and αA as follows.*

$$(A + B)(x) \equiv Ax + Bx, \ (\alpha A) x \equiv \alpha (Ax).$$

Theorem 11.4.4 *Let X and Y be finite dimensional inner product spaces. Then $\mathcal{L}(X, Y)$ is a vector space with the above definition of what it means to multiply by a scalar and add. Let $\{v_1, \cdots, v_n\}$ be an orthonormal basis for X and $\{w_1, \cdots, w_m\}$ be an orthonormal basis for Y. Then a basis for $\mathcal{L}(X, Y)$ is*

$$\{w_j \otimes v_i : i = 1, \cdots, n, \ j = 1, \cdots, m\}.$$

Proof: It is obvious that $\mathcal{L}(X, Y)$ is a vector space. It remains to verify the given set is a basis. Consider the following:

$$\left(\left(A - \sum_{k,l} (Av_k, w_l) w_l \otimes v_k\right) v_p, w_r\right) = (Av_p, w_r) -$$

$$\sum_{k,l} (Av_k, w_l)(v_p, v_k)(w_l, w_r)$$

$$= (Av_p, w_r) - \sum_{k,l} (Av_k, w_l) \delta_{pk} \delta_{rl} = (Av_p, w_r) - (Av_p, w_r) = 0.$$

Letting $A - \sum_{k,l} (Av_k, w_l) w_l \otimes v_k = B$, this shows that $Bv_p = 0$ since w_r is an arbitrary element of the basis for Y. Since v_p is an arbitrary element of the basis for X, it follows $B = 0$ as hoped. This has shown $\{w_j \otimes v_i : i = 1, \cdots, n, \ j = 1, \cdots, m\}$ spans $\mathcal{L}(X, Y)$.

It only remains to verify the $w_j \otimes v_i$ are linearly independent. Suppose then that

$$\sum_{i,j} c_{ij} w_j \otimes v_i = 0$$

Then do both sides to v_s. By definition this gives

$$0 = \sum_{i,j} c_{ij} w_j (v_s, v_i) = \sum_{i,j} c_{ij} w_j \delta_{si} = \sum_j c_{sj} w_j$$

Now the vectors $\{w_1, \cdots, w_m\}$ are independent because it is an orthonormal set and so the above requires $c_{sj} = 0$ for each j. Since s was arbitrary, this shows the linear transformations, $\{w_j \otimes v_i\}$ form a linearly independent set. ■

Note this shows the dimension of $\mathcal{L}(X, Y) = nm$. The theorem is also of enormous importance because it shows you can always consider an arbitrary linear transformation as a sum of rank one transformations whose properties are easily understood. The following theorem is also of great interest.

Theorem 11.4.5 *Let $A = \sum_{i,j} c_{ij} w_i \otimes v_j \in \mathcal{L}(X, Y)$ where as before, the vectors, $\{w_i\}$ are an orthonormal basis for Y and the vectors, $\{v_j\}$ are an orthonormal basis for X. Then if the matrix of A has entries M_{ij}, it follows that $M_{ij} = c_{ij}$.*

Proof: Recall

$$Av_i \equiv \sum_k M_{ki} w_k$$

Also

$$\begin{aligned} Av_i &= \sum_{k,j} c_{kj} w_k \otimes v_j (v_i) = \sum_{k,j} c_{kj} w_k (v_i, v_j) \\ &= \sum_{k,j} c_{kj} w_k \delta_{ij} = \sum_k c_{ki} w_k \end{aligned}$$

Therefore,

$$\sum_k M_{ki} w_k = \sum_k c_{ki} w_k$$

and so $M_{ki} = c_{ki}$ for all k. This happens for each i. ■

11.5 Least Squares

A common problem in experimental work is to find a straight line which approximates as well as possible a collection of points in the plane $\{(x_i, y_i)\}_{i=1}^p$. The usual way of dealing with these problems is by the method of least squares and it turns out that all these sorts of approximation problems can be reduced to $A\mathbf{x} = \mathbf{b}$ where the problem is to find the best $\mathbf{x}$ for solving this equation even when there is no solution.

Lemma 11.5.1 *Let V and W be finite dimensional inner product spaces and let $A : V \to W$ be linear. For each $y \in W$ there exists $x \in V$ such that*

$$|Ax - y| \leq |Ax_1 - y|$$

for all $x_1 \in V$. Also, $x \in V$ is a solution to this minimization problem if and only if x is a solution to the equation, $A^*Ax = A^*y$.

Proof: By Theorem 11.2.4 on Page 278 there exists a point, Ax_0, in the finite dimensional subspace, $A(V)$, of W such that for all $x \in V, |Ax - y|^2 \geq |Ax_0 - y|^2$. Also, from this theorem, this happens if and only if $Ax_0 - y$ is perpendicular to every $Ax \in A(V)$. Therefore, the solution is characterized by $(Ax_0 - y, Ax) = 0$ for all $x \in V$ which is the same as saying $(A^*Ax_0 - A^*y, x) = 0$ for all $x \in V$. In other words the solution is obtained by solving $A^*Ax_0 = A^*y$ for x_0. ■

Consider the problem of finding the least squares regression line in statistics. Suppose you have given points in the plane, $\{(x_i, y_i)\}_{i=1}^n$ and you would like to find constants m and b such that the line $y = mx + b$ goes through all these points. Of course this will be impossible in general. Therefore, try to find m, b such that you do the best you can to solve the system

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix}$$

which is of the form $\mathbf{y} = A\mathbf{x}$. In other words try to make $\left| A \begin{pmatrix} m \\ b \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right|^2$ as small as possible. According to what was just shown, it is desired to solve the following for m and b.

$$A^* A \begin{pmatrix} m \\ b \end{pmatrix} = A^* \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Since $A^* = A^T$ in this case,

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

Solving this system of equations for m and b,

$$m = \frac{-\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right) + \left(\sum_{i=1}^n x_i y_i\right) n}{\left(\sum_{i=1}^n x_i^2\right) n - \left(\sum_{i=1}^n x_i\right)^2}$$

and

$$b = \frac{-\left(\sum_{i=1}^n x_i\right)\sum_{i=1}^n x_i y_i + \left(\sum_{i=1}^n y_i\right)\sum_{i=1}^n x_i^2}{\left(\sum_{i=1}^n x_i^2\right) n - \left(\sum_{i=1}^n x_i\right)^2}.$$

One could clearly do a least squares fit for curves of the form $y = ax^2 + bx + c$ in the same way. In this case you solve as well as possible for a, b, and c the system

$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

using the same techniques.

11.6 Fredholm Alternative Again

The best context in which to study the Fredholm alternative is in inner product spaces. This is done here.

Definition 11.6.1 *Let S be a subset of an inner product space, X. Define*

$$S^{\perp} \equiv \{x \in X : (x, s) = 0 \text{ for all } s \in S\}.$$

The following theorem also follows from the above lemma. It is sometimes called the Fredholm alternative.

Theorem 11.6.2 *Let $A : V \to W$ where A is linear and V and W are inner product spaces. Then $A(V) = \ker(A^*)^{\perp}$.*

Proof: Let $y = Ax$ so $y \in A(V)$. Then if $A^*z = 0$,

$$(y, z) = (Ax, z) = (x, A^*z) = 0$$

showing that $y \in \ker(A^*)^{\perp}$. Thus $A(V) \subseteq \ker(A^*)^{\perp}$.

Now suppose $y \in \ker(A^*)^{\perp}$. Does there exists x such that $Ax = y$? Since this might not be immediately clear, take the least squares solution to the problem. Thus let x be a solution to $A^*Ax = A^*y$. It follows $A^*(y - Ax) = 0$ and so $y - Ax \in \ker(A^*)$ which implies from the assumption about y that $(y - Ax, y) = 0$. Also, since Ax is the closest point to y in $A(V)$, Theorem 11.2.4 on Page 278 implies that $(y - Ax, Ax_1) = 0$ for all $x_1 \in V$. In particular this is true for $x_1 = x$ and so $0 = (y - Ax, y) - \overbrace{(y - Ax, Ax)}^{=0} = |y - Ax|^2$, showing that $y = Ax$. Thus $A(V) \supseteq \ker(A^*)^{\perp}$. ■

Corollary 11.6.3 *Let $A, V,$ and W be as described above. If the only solution to $A^*y = 0$ is $y = 0$, then A is onto W.*

Proof: If the only solution to $A^*y = 0$ is $y = 0$, then $\ker(A^*) = \{0\}$ and so every vector from W is contained in $\ker(A^*)^{\perp}$ and by the above theorem, this shows $A(V) = W$. ■

11.7 Exercises

1. Find the best solution to the system

$$\begin{aligned} x + 2y &= 6 \\ 2x - y &= 5 \\ 3x + 2y &= 0 \end{aligned}$$

2. Find an orthonormal basis for $\mathbb{R}^3$, $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ given that $\mathbf{w}_1$ is a multiple of the vector $(1, 1, 2)$.

3. Suppose $A = A^T$ is a symmetric real $n \times n$ matrix which has all positive eigenvalues. Define

$$(\mathbf{x}, \mathbf{y}) \equiv (A\mathbf{x}, \mathbf{y}).$$

Show this is an inner product on $\mathbb{R}^n$. What does the Cauchy Schwarz inequality say in this case?

4. Let $||\mathbf{x}||_\infty \equiv \max\{|x_j| : j = 1, 2, \cdots, n\}$. Show this is a norm on $\mathbb{C}^n$. Here

$$\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T.$$

Show

$$||\mathbf{x}||_\infty \leq |\mathbf{x}| \equiv (\mathbf{x}, \mathbf{x})^{1/2}$$

where the above is the usual inner product on $\mathbb{C}^n$.

5. Let $||\mathbf{x}||_1 \equiv \sum_{j=1}^n |x_j|$ .Show this is a norm on $\mathbb{C}^n$. Here $\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T$. Show

$$||\mathbf{x}||_1 \geq |\mathbf{x}| \equiv (\mathbf{x}, \mathbf{x})^{1/2}$$

where the above is the usual inner product on $\mathbb{C}^n$. Show there cannot exist an inner product such that this norm comes from the inner product as described above for inner product spaces.

6. Show that if $||\cdot||$ is any norm on any vector space, then $|||x|| - ||y||| \leq ||x - y||$.

7. Relax the assumptions in the axioms for the inner product. Change the axiom about $(x, x) \geq 0$ and equals 0 if and only if $x = 0$ to simply read $(x, x) \geq 0$. Show the Cauchy Schwarz inequality still holds in the following form. $|(x, y)| \leq (x, x)^{1/2} (y, y)^{1/2}$.

8. Let H be an inner product space and let $\{u_k\}_{k=1}^n$ be an orthonormal basis for H. Show

$$(x, y) = \sum_{k=1}^n (x, u_k) \overline{(y, u_k)}.$$

9. Let the vector space V consist of real polynomials of degree no larger than 3. Thus a typical vector is a polynomial of the form $a + bx + cx^2 + dx^3$. For $p, q \in V$ define the inner product, $(p, q) \equiv \int_0^1 p(x) q(x) dx$. Show this is indeed an inner product. Then state the Cauchy Schwarz inequality in terms of this inner product. Show $\{1, x, x^2, x^3\}$ is a basis for V. Finally, find an orthonormal basis for V. This is an example of some orthonormal polynomials.

10. Let P_n denote the polynomials of degree no larger than $n - 1$ which are defined on an interval $[a, b]$. Let $\{x_1, \cdots, x_n\}$ be n distinct points in $[a, b]$. Now define for $p, q \in P_n$,

$$(p, q) \equiv \sum_{j=1}^n p(x_j) \overline{q(x_j)}$$

Show this yields an inner product on P_n. **Hint:** Most of the axioms are obvious. The one which says $(p, p) = 0$ if and only if $p = 0$ is the only interesting one. To verify this one, note that a nonzero polynomial of degree no more than $n - 1$ has at most $n - 1$ zeros.

11. Let $C([0, 1])$ denote the vector space of continuous real valued functions defined on $[0, 1]$. Let the inner product be given as

$$(f, g) \equiv \int_0^1 f(x) g(x) dx$$

Show this is an inner product. Also let V be the subspace described in Problem 9. Using the result of this problem, find the vector in V which is closest to x^4.

12. A **regular Sturm Liouville problem** involves the differential equation, for an unknown function of x which is denoted here by y,

$$(p(x) y')' + (\lambda q(x) + r(x)) y = 0, \ x \in [a, b]$$

and it is assumed that $p(t), q(t) > 0$ for any $t \in [a,b]$ and also there are boundary conditions,

$$\begin{aligned} C_1 y(a) + C_2 y'(a) &= 0 \\ C_3 y(b) + C_4 y'(b) &= 0 \end{aligned}$$

where

$$C_1^2 + C_2^2 > 0, \text{ and } C_3^2 + C_4^2 > 0.$$

There is an immense theory connected to these important problems. The constant, λ is called an eigenvalue. Show that if y is a solution to the above problem corresponding to $\lambda = \lambda_1$ and if z is a solution corresponding to $\lambda = \lambda_2 \neq \lambda_1$, then

$$\int_a^b q(x) y(x) z(x)\, dx = 0. \tag{11.9}$$

and this defines an inner product. **Hint:** Do something like this:

$$\begin{aligned} (p(x) y')' z + (\lambda_1 q(x) + r(x)) yz &= 0, \\ (p(x) z')' y + (\lambda_2 q(x) + r(x)) zy &= 0. \end{aligned}$$

Now subtract and either use integration by parts or show

$$(p(x) y')' z - (p(x) z')' y = ((p(x) y') z - (p(x) z') y)'$$

and then integrate. Use the boundary conditions to show that $y'(a) z(a) - z'(a) y(a) = 0$ and $y'(b) z(b) - z'(b) y(b) = 0$. The formula, 11.9 is called an orthogonality relation. It turns out there are typically infinitely many eigenvalues and it is interesting to write given functions as an infinite series of these "eigenfunctions".

13. Consider the continuous functions defined on $[0,\pi]$, $C([0,\pi])$. Show $(f,g) \equiv \int_0^\pi fg dx$ is an inner product on this vector space. Show the functions $\left\{\sqrt{\frac{2}{\pi}}\sin(nx)\right\}_{n=1}^\infty$ are an orthonormal set. What does this mean about the dimension of the vector space $C([0,\pi])$? Now let $V_N = \text{span}\left(\sqrt{\frac{2}{\pi}}\sin(x), \cdots, \sqrt{\frac{2}{\pi}}\sin(Nx)\right)$. For $f \in C([0,\pi])$ find a formula for the vector in V_N which is closest to f with respect to the norm determined from the above inner product. This is called the N^{th} partial sum of the Fourier series of f. An important problem is to determine whether and in what way this Fourier series converges to the function f. The norm which comes from this inner product is sometimes called the mean square norm.

14. Consider the subspace $V \equiv \ker(A)$ where

$$A = \begin{pmatrix} 1 & 4 & -1 & -1 \\ 2 & 1 & 2 & 3 \\ 4 & 9 & 0 & 1 \\ 5 & 6 & 3 & 4 \end{pmatrix}$$

Find an orthonormal basis for V. **Hint:** You might first find a basis and then use the Gram Schmidt procedure.

15. The Gram Schmidt process starts with a basis for a subspace $\{v_1, \cdots, v_n\}$ and produces an orthonormal basis for the same subspace $\{u_1, \cdots, u_n\}$ such that

$$\text{span}(v_1, \cdots, v_k) = \text{span}(u_1, \cdots, u_k)$$

for each k. Show that in the case of $\mathbb{R}^m$ the QR factorization does the same thing. More specifically, if

$$A = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$$

and if

$$A = QR \equiv \begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{pmatrix} R$$

then the vectors $\{\mathbf{q}_1, \cdots, \mathbf{q}_n\}$ is an orthonormal set of vectors and for each k,

$$\text{span}(\mathbf{q}_1, \cdots, \mathbf{q}_k) = \text{span}(\mathbf{v}_1, \cdots, \mathbf{v}_k)$$

16. Verify the parallelogram identify for any inner product space,

$$|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2.$$

Why is it called the parallelogram identity?

17. Let H be an inner product space and let $K \subseteq H$ be a nonempty convex subset. This means that if $k_1, k_2 \in K$, then the line segment consisting of points of the form

$$tk_1 + (1-t)k_2 \text{ for } t \in [0,1]$$

is also contained in K. Suppose for each $x \in H$, there exists Px defined to be a point of K closest to x. Show that Px is unique so that P actually is a map. **Hint:** Suppose z_1 and z_2 both work as closest points. Consider the midpoint, $(z_1+z_2)/2$ and use the parallelogram identity of Problem 16 in an auspicious manner.

18. In the situation of Problem 17 suppose K is a closed convex subset and that H is complete. This means every Cauchy sequence converges. Recall from calculus a sequence $\{k_n\}$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists N_ε such that whenever $m, n > N_\varepsilon$, it follows $|k_m - k_n| < \varepsilon$. Let $\{k_n\}$ be a sequence of points of K such that

$$\lim_{n\to\infty} |x - k_n| = \inf\{|x-k| : k \in K\}$$

This is called a minimizing sequence. Show there exists a unique $k \in K$ such that $\lim_{n\to\infty} |k_n - k|$ and that $k = Px$. That is, there exists a well defined projection map onto the convex subset of H. **Hint:** Use the parallelogram identity in an auspicious manner to show $\{k_n\}$ is a Cauchy sequence which must therefore converge. Since K is closed it follows this will converge to something in K which is the desired vector.

19. Let H be an inner product space which is also complete and let P denote the projection map onto a convex closed subset, K. Show this projection map is characterized by the inequality
$$\text{Re}\,(k - Px, x - Px) \leq 0$$
for all $k \in K$. That is, a point $z \in K$ equals Px if and only if the above variational inequality holds. This is what that inequality is called. This is because k is allowed to vary and the inequality continues to hold for all $k \in K$.

20. Using Problem 19 and Problems 17 - 18 show the projection map, P onto a closed convex subset is Lipschitz continuous with Lipschitz constant 1. That is
$$|Px - Py| \leq |x - y|$$

21. Give an example of two vectors in $\mathbb{R}^4$ or $\mathbb{R}^3$ $\mathbf{x}, \mathbf{y}$ and a subspace V such that $\mathbf{x} \cdot \mathbf{y} = 0$ but $P\mathbf{x} \cdot P\mathbf{y} \neq 0$ where P denotes the projection map which sends $\mathbf{x}$ to its closest point on V.

22. Suppose you are given the data, $(1, 2), (2, 4), (3, 8), (0, 0)$. Find the linear regression line using the formulas derived above. Then graph the given data along with your regression line.

23. Generalize the least squares procedure to the situation in which data is given and you desire to fit it with an expression of the form $y = af(x) + bg(x) + c$ where the problem would be to find a, b and c in order to minimize the error. Could this be generalized to higher dimensions? How about more functions?

24. Let $A \in \mathcal{L}(X, Y)$ where X and Y are finite dimensional vector spaces with the dimension of X equal to n. Define $\text{rank}\,(A) \equiv \dim (A(X))$ and $\text{nullity}(A) \equiv \dim(\ker(A))$. Show that $\text{nullity}(A) + \text{rank}\,(A) = \dim(X)$. **Hint:** Let $\{x_i\}_{i=1}^r$ be a basis for $\ker(A)$ and let $\{x_i\}_{i=1}^r \cup \{y_i\}_{i=1}^{n-r}$ be a basis for X. Then show that $\{Ay_i\}_{i=1}^{n-r}$ is linearly independent and spans AX.

25. Let A be an $m \times n$ matrix. Show the column rank of A equals the column rank of A^*A. Next verify column rank of A^*A is no larger than column rank of A^*. Next justify the following inequality to conclude the column rank of A equals the column rank of A^*.
$$\text{rank}\ (A) = \text{rank}\ (A^*A) \leq \text{rank}\ (A^*) \leq$$
$$= \text{rank}\ (AA^*) \leq \text{rank}\ (A).$$
Hint: Start with an orthonormal basis, $\{A\mathbf{x}_j\}_{j=1}^r$ of $A(\mathbb{F}^n)$ and verify $\{A^*A\mathbf{x}_j\}_{j=1}^r$ is a basis for $A^*A(\mathbb{F}^n)$.

26. Let A be a real $m \times n$ matrix and let $A = QR$ be the QR factorization with Q orthogonal and R upper triangular. Show that there exists a solution $\mathbf{x}$ to the equation
$$R^T R\mathbf{x} = R^T Q^T \mathbf{b}$$
and that this solution is also a least squares solution defined above such that $A^T A\mathbf{x} = A^T\mathbf{b}$.

11.8 The Determinant And Volume

The determinant is the essential algebraic tool which provides a way to give a unified treatment of the concept of p dimensional volume of a parallelepiped in $\mathbb{R}^M$. Here is the definition of what is meant by such a thing.

Definition 11.8.1 *Let $\mathbf{u}_1, \cdots, \mathbf{u}_p$ be vectors in $\mathbb{R}^M, M \geq p$. The parallelepiped determined by these vectors will be denoted by $P(\mathbf{u}_1, \cdots, \mathbf{u}_p)$ and it is defined as*

$$P(\mathbf{u}_1, \cdots, \mathbf{u}_p) \equiv \left\{ \sum_{j=1}^{p} s_j \mathbf{u}_j : s_j \in [0,1] \right\} = UQ,\ \ Q = [0,1]^p$$

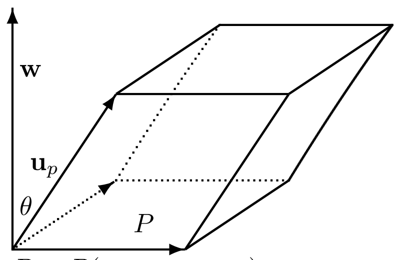
where $U = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_p \end{pmatrix}$. The volume of this parallelepiped is defined as

$$\text{volume of } P(\mathbf{u}_1, \cdots, \mathbf{u}_p) \equiv v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p)) \equiv (\det(G))^{1/2}.$$

where $G_{ij} = \mathbf{u}_i \cdot \mathbf{u}_j$. This $G = U^T U$ is called the metric tensor. If the vectors $\mathbf{u}_i$ are dependent, this definition will give the volume to be 0.

First lets observe the last assertion is true. Say $\mathbf{u}_i = \sum_{j \neq i} \alpha_j \mathbf{u}_j$. Then the i^{th} row of G is a linear combination of the other rows using the scalars α_j and so from the properties of the determinant, the determinant of this matrix is indeed zero as it should be. Indeed, $\mathbf{u}_i \cdot \mathbf{u}_k = \sum_{j \neq i} \alpha_j \mathbf{u}_j \cdot \mathbf{u}_k$.

A parallelepiped is a sort of a squashed box. Here is a picture which shows the relationship between $P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})$ and $P(\mathbf{u}_1, \cdots, \mathbf{u}_p)$. In a sense, we can define the volume any way desired, but if it is to be reasonable, the following relationship must hold. The appropriate definition of the volume of $P(\mathbf{u}_1, \cdots, \mathbf{u}_p)$ in terms of $P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})$ is $v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p)) =$



$P = P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})$

$$|\mathbf{u}_p \cdot \mathbf{w}| \, v(P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})) \tag{11.10}$$

where $\mathbf{w}$ is any unit vector perpendicular to each of $\mathbf{u}_1, \cdots, \mathbf{u}_{p-1}$. Note $|\mathbf{u}_p \cdot \mathbf{w}| = |\mathbf{u}_p| |\cos\theta|$ from the geometric meaning of the dot product. In the case where $p = 1$, the parallelepiped $P(\mathbf{v})$ consists of the single vector and the one dimensional volume should be $|\mathbf{v}| = \left(\mathbf{v}^T \mathbf{v}\right)^{1/2} = (\mathbf{v} \cdot \mathbf{v})^{1/2}$. Now having made this definition, I will show that $\det(G)^{1/2}$ is the appropriate definition of $v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p))$ for every p.

As just pointed out, this is the only reasonable definition of volume in the case of one vector. The next theorem shows that it is the only reasonable definition of volume of a parallelepiped in the case of p vectors because 11.10 holds.

Theorem 11.8.2 *If we desire 11.10 to hold for any $\mathbf{w}$ perpendicular to each $\mathbf{u}_i$, then we obtain the definition of 11.8.1 for $v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p))$ in terms of determinants.*

Proof: So assume we want 11.10 to hold. Suppose the determinant formula holds for $P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})$. It is necessary to show that if $\mathbf{w}$ is a unit vector perpendicular to each $\mathbf{u}_1, \cdots, \mathbf{u}_{p-1}$ then $|\mathbf{u}_p \cdot \mathbf{w}| \, v(P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1}))$ reduces to $\det(G)^{1/2}$. By the Gram Schmidt procedure there is $(\mathbf{w}_1, \cdots, \mathbf{w}_p)$ an orthonormal basis for $\text{span}(\mathbf{u}_1, \cdots, \mathbf{u}_p)$ such that $\text{span}(\mathbf{w}_1, \cdots, \mathbf{w}_k) = \text{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k)$ for each $k \leq p$. We can pick $\mathbf{w}_p = \mathbf{w}$ the given unit vector perpendicular to each $\mathbf{u}_i$. First note that since $\{\mathbf{w}_k\}_{k=1}^p$ is an orthonormal basis for $\text{span}(\mathbf{u}_1, \cdots, \mathbf{u}_p)$,

$$\mathbf{u}_j = \sum_{k=1}^{p} (\mathbf{u}_j \cdot \mathbf{w}_k) \mathbf{w}_k, \quad \mathbf{u}_j \cdot \mathbf{u}_i = \sum_{k=1}^{p} (\mathbf{u}_j \cdot \mathbf{w}_k)(\mathbf{u}_i \cdot \mathbf{w}_k)$$

Therefore, the ij^{th} entry of the $p \times p$ matrix $U^T U$ is just

$$\left(U^T U\right)_{ij} = \sum_{r=1}^{p} (\mathbf{u}_i \cdot \mathbf{w}_r)(\mathbf{w}_r \cdot \mathbf{u}_j)$$

which is the product of a $p \times p$ matrix M whose rj^{th} entry is $\mathbf{w}_r \cdot \mathbf{u}_j$ with its transpose. The vector $\mathbf{w}_p$ is a unit vector perpendicular to each $\mathbf{u}_j$ for $j \leq p - 1$ so $\mathbf{w}_p \cdot \mathbf{u}_j = 0$ if $j < p$.

Now consider the vector

$$\mathbf{N} \equiv \det \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_{p-1} & \mathbf{w}_p \\ \mathbf{u}_1 \cdot \mathbf{w}_1 & \cdots & \mathbf{u}_1 \cdot \mathbf{w}_{p-1} & \overset{=0}{\mathbf{u}_1 \cdot \mathbf{w}_p} \\ \vdots & & \vdots & \vdots \\ \mathbf{u}_{p-1} \cdot \mathbf{w}_1 & \cdots & \mathbf{u}_{p-1} \cdot \mathbf{w}_{p-1} & \overset{=0}{\mathbf{u}_{p-1} \cdot \mathbf{w}_p} \end{pmatrix}$$

which results from formally expanding along the top row. Note you would get the same thing expanding along the last column because as just noted, the last column on the right is 0 except for the top entry, so every cofactor A_{1k} for the $1k^{th}$ position is $\pm$ a determinant which has a column of zeros. Thus $\mathbf{N}$ is a multiple of $\mathbf{w}_p$. Hence, for $j < p, \mathbf{N} \cdot \mathbf{u}_j = 0$. From what was just discussed and induction, $v\left(P\left(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1}\right)\right) = \pm A_{1p} = \mathbf{N} \cdot \mathbf{w}_p$. Also $\mathbf{N} \cdot \mathbf{u}_p$ equals

$$\det \begin{pmatrix} \mathbf{u}_p \cdot \mathbf{w}_1 & \cdots & \mathbf{u}_p \cdot \mathbf{w}_{p-1} & \mathbf{u}_p \cdot \mathbf{w}_p \\ \mathbf{u}_1 \cdot \mathbf{w}_1 & \cdots & \mathbf{u}_1 \cdot \mathbf{w}_{p-1} & \overset{=0}{\mathbf{u}_1 \cdot \mathbf{w}_p} \\ \vdots & & \vdots & \vdots \\ \mathbf{u}_{p-1} \cdot \mathbf{w}_1 & \cdots & \mathbf{u}_{p-1} \cdot \mathbf{w}_{p-1} & \overset{=0}{\mathbf{u}_{p-1} \cdot \mathbf{w}_p} \end{pmatrix} = \pm \det(M)$$

Thus from induction and expanding along the last column,

$$\begin{aligned} \left|\mathbf{u}_p \cdot \mathbf{w}_p\right| v\left(P\left(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1}\right)\right) &= \left|\mathbf{N} \cdot \mathbf{u}_p\right| = \det\left(M^T M\right)^{1/2} \\ &= \det\left(U^T U\right)^{1/2} = \det(G)^{1/2}. \end{aligned}$$

Now $\mathbf{w}_p = \mathbf{w}$ the unit vector perpendicular to each $\mathbf{u}_j$ for $j \leq p-1$. Thus if 11.10, then the claimed determinant identity holds. ■

The theorem shows that the only reasonable definition of p dimensional volume of a parallelepiped is the one given in the above definition. Recall that these vectors are in $\mathbb{R}^M$. What is the role of $\mathbb{R}^M$? It is just to provide an inner product. That is its only function. If $p = M$, then $\det\left(U^T U\right) = \det\left(U^T\right)\det(U) = \det(U)^2$ and so $\det(G)^{1/2} = |\det(U)|$.

11.9 Exercises

1. Here are three vectors in $\mathbb{R}^4 : (1,2,0,3)^T, (2,1,-3,2)^T, (0,0,1,2)^T$. Find the three dimensional volume of the parallelepiped determined by these three vectors.

2. Here are two vectors in $\mathbb{R}^4 : (1,2,0,3)^T, (2,1,-3,2)^T$. Find the volume of the parallelepiped determined by these two vectors.

3. Here are three vectors in $\mathbb{R}^2 : (1,2)^T, (2,1)^T, (0,1)^T$. Find the three dimensional volume of the parallelepiped determined by these three vectors. Recall that from the above theorem, this should equal 0.

4. Find the equation of the plane through the three points $(1,2,3), (2,-3,1), (1,1,7)$.

5. Let T map a vector space V to itself. Explain why T is one to one if and only if T is onto. It is in the text, but do it again in your own words.

6. ↑Let all matrices be complex with complex field of scalars and let A be an $n \times n$ matrix and B a $m \times m$ matrix while X will be an $n \times m$ matrix. The problem is to consider solutions to Sylvester's equation. Solve the following equation for X

$$AX - XB = C$$

where C is an arbitrary $n \times m$ matrix. Show there exists a unique solution if and only if $\sigma(A) \cap \sigma(B) = \emptyset$. **Hint:** If $q(\lambda)$ is a polynomial, show first that if $AX - XB = 0$, then $q(A)X - Xq(B) = 0$. Next define the linear map T which maps the $n \times m$ matrices to the $n \times m$ matrices as follows.

$$TX \equiv AX - XB$$

Show that the only solution to $TX = 0$ is $X = 0$ so that T is one to one if and only if $\sigma(A) \cap \sigma(B) = \emptyset$. Do this by using the first part for $q(\lambda)$ the characteristic polynomial for B and then use the Cayley Hamilton theorem. Explain why $q(A)^{-1}$ exists if and only if the condition $\sigma(A) \cap \sigma(B) = \emptyset$.

7. Recall the Binet Cauchy theorem, Theorem 3.3.14. What is the geometric meaning of the Binet Cauchy theorem?

8. For W a subspace of V, W is said to have a complementary subspace [15] W' if $W \oplus W' = V$. Suppose that both W, W' are invariant with respect to $A \in \mathcal{L}(V,V)$. Show that for any polynomial $f(\lambda)$, if $f(A)x \in W$, then there exists $w \in W$ such that $f(A)x = f(A)w$. A subspace W is called A admissible if it is A invariant and the condition of this problem holds.

9. ↑ Return to Theorem 9.3.5 about the existence of a basis $\beta = \left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$ for V where $A \in \mathcal{L}(V,V)$. Adapt the statement and proof to show that if W is A admissible, then it has a complementary subspace which is also A invariant. **Hint:**

   The modified version of the theorem is: Suppose $A \in \mathcal{L}(V,V)$ and the minimal polynomial of A is $\phi(\lambda)^m$ where $\phi(\lambda)$ is a monic irreducible polynomial. Also suppose that W is an A admissible subspace. Then there exists a basis for V which is of the form $\beta = \left\{\beta_{x_1}, \cdots, \beta_{x_p}, v_1, \cdots, v_m\right\}$ where $\{v_1, \cdots, v_m\}$ is a basis of W. Thus $\text{span}\left(\beta_{x_1}, \cdots, \beta_{x_p}\right)$ is the A invariant complementary subspace for W. You may want to use the fact that $\phi(A)(V) \cap W = \phi(A)(W)$ which follows easily because W is A admissible. Then use this fact to show that $\phi(A)(W)$ is also A admissible.

10. Let U, H be finite dimensional inner product spaces. (More generally, complete inner product spaces.) Let A be a linear map from U to H. Thus AU is a subspace of H. For $\mathbf{g} \in AU$, define $A^{-1}\mathbf{g}$ to be the unique element of $\{\mathbf{x} : A\mathbf{x} = \mathbf{g}\}$ which is closest to $\mathbf{0}$. Then define $(\mathbf{h}, \mathbf{g})_{AU} \equiv \left(A^{-1}\mathbf{g}, A^{-1}\mathbf{h}\right)_U$. Show that this is a well defined inner product. Let U, H be finite dimensional inner product spaces. (More generally, complete inner product spaces.) Let A be a linear map from U to H. Thus AU is a subspace of H. For $\mathbf{g} \in AU$, define $A^{-1}\mathbf{g}$ to be the unique element of $\{\mathbf{x} : A\mathbf{x} = \mathbf{g}\}$ which is closest to $\mathbf{0}$. Then define $(\mathbf{h}, \mathbf{g})_{AU} \equiv \left(A^{-1}\mathbf{g}, A^{-1}\mathbf{h}\right)_U$. Show that this is a well defined inner product and that if A is one to one, then $\|\mathbf{h}\|_{AU} = \left\|A^{-1}\mathbf{h}\right\|_U$ and $\|A\mathbf{x}\|_{AU} = \|\mathbf{x}\|_U$.

Chapter 12

Self Adjoint Operators

12.1 Simultaneous Diagonalization

Recall the following definition of what it means for a matrix to be diagonalizable.

Definition 12.1.1 *Let A be an $n \times n$ matrix. It is said to be diagonalizable if there exists an invertible matrix S such that*

$$S^{-1}AS = D$$

where D is a diagonal matrix.

Also, here is a useful observation.

Observation 12.1.2 *If A is an $n \times n$ matrix and $AS = SD$ for D a diagonal matrix, then each column of S is an eigenvector or else it is the zero vector. This follows from observing that for $\mathbf{s}_k$ the k^{th} column of S and from the way we multiply matrices,*

$$A\mathbf{s}_k = \lambda_k \mathbf{s}_k$$

It is sometimes interesting to consider the problem of finding a single similarity transformation which will diagonalize all the matrices in some set.

Lemma 12.1.3 *Let A be an $n \times n$ matrix and let B be an $m \times m$ matrix. Denote by C the matrix*

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

Then C is diagonalizable if and only if both A and B are diagonalizable.

Proof: Suppose $S_A^{-1}AS_A = D_A$ and $S_B^{-1}BS_B = D_B$ where D_A and D_B are diagonal matrices. You should use block multiplication to verify that $S \equiv \begin{pmatrix} S_A & 0 \\ 0 & S_B \end{pmatrix}$ is such that $S^{-1}CS = D_C$, a diagonal matrix.

Conversely, suppose C is diagonalized by $S = (\mathbf{s}_1, \cdots, \mathbf{s}_{n+m})$. Thus S has columns $\mathbf{s}_i$. For each of these columns, write in the form

$$\mathbf{s}_i = \begin{pmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{pmatrix}$$

where $\mathbf{x}_i \in \mathbb{F}^n$ and where $\mathbf{y}_i \in \mathbb{F}^m$. The result is

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

where S_{11} is an $n \times n$ matrix and S_{22} is an $m \times m$ matrix. Then there is a diagonal matrix, D_1 being $n \times n$ and D_2 $m \times m$ such that

$$D = \text{diag}(\lambda_1, \cdots, \lambda_{n+m}) = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}$$

such that

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$
$$= \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}$$

Hence by block multiplication

$$AS_{11} = S_{11}D_1,\ BS_{22} = S_{22}D_2$$

$$BS_{21} = S_{21}D_1,\ AS_{12} = S_{12}D_2$$

It follows each of the $\mathbf{x}_i$ is an eigenvector of A or else is the zero vector and that each of the $\mathbf{y}_i$ is an eigenvector of B or is the zero vector. If there are n linearly independent $\mathbf{x}_i$, then A is diagonalizable by Theorem 8.3.12 on Page 8.3.12.

The row rank of the matrix $(\mathbf{x}_1, \cdots, \mathbf{x}_{n+m})$ must be n because if this is not so, the rank of S would be less than $n+m$ which would mean S^{-1} does not exist. Therefore, since the column rank equals the row rank, this matrix has column rank equal to n and this means there are n linearly independent eigenvectors of A implying that A is diagonalizable. Similar reasoning applies to B. ■

The following corollary follows from the same type of argument as the above.

Corollary 12.1.4 *Let A_k be an $n_k \times n_k$ matrix and let C denote the block diagonal*

$$\left(\sum_{k=1}^{r} n_k\right) \times \left(\sum_{k=1}^{r} n_k\right)$$

matrix given below.

$$C \equiv \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}.$$

Then C is diagonalizable if and only if each A_k is diagonalizable.

Definition 12.1.5 *A set, $\mathcal{F}$ of $n \times n$ matrices is said to be simultaneously diagonalizable if and only if there exists a single invertible matrix S such that for every $A \in \mathcal{F}$, $S^{-1}AS = D_A$ where D_A is a diagonal matrix. $\mathcal{F}$ is a commuting family of matrices if whenever $A, B \in \mathcal{F}$, $AB = BA$.*

Lemma 12.1.6 *If $\mathcal{F}$ is a set of $n \times n$ matrices which is simultaneously diagonalizable, then $\mathcal{F}$ is a commuting family of matrices.*

Proof: Let $A, B \in \mathcal{F}$ and let S be a matrix which has the property that $S^{-1}AS$ is a diagonal matrix for all $A \in \mathcal{F}$. Then $S^{-1}AS = D_A$ and $S^{-1}BS = D_B$ where D_A and D_B are diagonal matrices. Since diagonal matrices commute,

$$\begin{aligned} AB &= SD_AS^{-1}SD_BS^{-1} = SD_AD_BS^{-1} \\ &= SD_BD_AS^{-1} = SD_BS^{-1}SD_AS^{-1} = BA. \end{aligned}$$

Lemma 12.1.7 *Let D be a diagonal matrix of the form*

$$D \equiv \begin{pmatrix} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r I_{n_r} \end{pmatrix}, \tag{12.1}$$

where I_{n_i} denotes the $n_i \times n_i$ identity matrix and $\lambda_i \neq \lambda_j$ for $i \neq j$ and suppose B is a matrix which commutes with D. Then B is a block diagonal matrix of the form

$$B = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_r \end{pmatrix} \tag{12.2}$$

where B_i is an $n_i \times n_i$ matrix.

Proof: Let $B = (B_{ij})$ where $B_{ii} = B_i$ a block matrix as above in 12.2.

$$\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22} & \ddots & B_{2r} \\ \vdots & \ddots & \ddots & \vdots \\ B_{r1} & B_{r2} & \cdots & B_{rr} \end{pmatrix}$$

Then by block multiplication, since B is given to commute with D,

$$\lambda_j B_{ij} = \lambda_i B_{ij}$$

Therefore, if $i \neq j$, $B_{ij} = 0$. ■

Lemma 12.1.8 *Let $\mathcal{F}$ denote a commuting family of $n \times n$ matrices such that each $A \in \mathcal{F}$ is diagonalizable. Then $\mathcal{F}$ is simultaneously diagonalizable.*

Proof: First note that if every matrix in $\mathcal{F}$ has only one eigenvalue, there is nothing to prove. This is because for A such a matrix,

$$S^{-1}AS = \lambda I$$

and so

$$A = \lambda I$$

Thus all the matrices in $\mathcal{F}$ are diagonal matrices and you could pick any S to diagonalize them all. Therefore, without loss of generality, assume some matrix in $\mathcal{F}$ has more than one eigenvalue.

The significant part of the lemma is proved by induction on n. If $n = 1$, there is nothing to prove because all the 1×1 matrices are already diagonal matrices. Suppose then that the theorem is true for all $k \leq n-1$ where $n \geq 2$ and let $\mathcal{F}$ be a commuting family of diagonalizable $n \times n$ matrices. Pick $A \in \mathcal{F}$ which has more than one eigenvalue and let S be an invertible matrix such that $S^{-1}AS = D$ where D is of the form given in 12.1. By permuting the columns of S there is no loss of generality in assuming D has this form. Now denote by $\widetilde{\mathcal{F}}$ the collection of matrices, $\left\{S^{-1}CS : C \in \mathcal{F}\right\}$. Note $\widetilde{\mathcal{F}}$ features the single matrix S.

It follows easily that $\widetilde{\mathcal{F}}$ is also a commuting family of diagonalizable matrices. By Lemma 12.1.7 every $B \in \widetilde{\mathcal{F}}$ is a block diagonal matrix of the form given in 12.2 because each of these commutes with D described above as $S^{-1}AS$ and so by block multiplication, the diagonal blocks B_i corresponding to different $B \in \widetilde{\mathcal{F}}$ commute.

By Corollary 12.1.4 each of these blocks is diagonalizable. This is because B is known to be so. Therefore, by induction, since all the blocks are no larger than $n-1 \times n-1$ thanks to the assumption that A has more than one eigenvalue, there exist invertible $n_i \times n_i$ matrices, T_i such that $T_i^{-1}B_iT_i$ is a diagonal matrix whenever B_i is one of the matrices making up the block diagonal of any $B \in \widetilde{\mathcal{F}}$. It follows that for T defined by

$$T \equiv \begin{pmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & T_r \end{pmatrix},$$

then $T^{-1}BT =$ a diagonal matrix for every $B \in \widetilde{\mathcal{F}}$ including D. Consider ST. It follows that for all $C \in \mathcal{F}$,

$$T^{-1} \overbrace{S^{-1}CS}^{\text{something in } \widetilde{\mathcal{F}}} T = (ST)^{-1} C (ST) = \text{ a diagonal matrix.} \blacksquare$$

Theorem 12.1.9 *Let $\mathcal{F}$ denote a family of matrices which are diagonalizable. Then $\mathcal{F}$ is simultaneously diagonalizable if and only if $\mathcal{F}$ is a commuting family.*

Proof: If $\mathcal{F}$ is a commuting family, it follows from Lemma 12.1.8 that it is simultaneously diagonalizable. If it is simultaneously diagonalizable, then it follows from Lemma 12.1.6 that it is a commuting family. ■

12.2 Schur's Theorem

Recall that for a linear transformation, $L \in \mathcal{L}(V,V)$ for V a finite dimensional inner product space, it could be represented in the form

$$L = \sum_{ij} l_{ij} \mathbf{v}_i \otimes \mathbf{v}_j$$

where $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is an orthonormal basis. Of course different bases will yield different matrices, (l_{ij}). Schur's theorem gives the existence of a basis in an inner product space such that (l_{ij}) is particularly simple.

Definition 12.2.1 *Let $L \in \mathcal{L}(V, V)$ where V is a vector space. Then a subspace U of V is L invariant if $L(U) \subseteq U$.*

In what follows, $\mathbb{F}$ will be the field of scalars, usually $\mathbb{C}$ but maybe $\mathbb{R}$.

Theorem 12.2.2 *Let $L \in \mathcal{L}(H, H)$ for H a finite dimensional inner product space such that the restriction of L^* to every L invariant subspace has its eigenvalues in $\mathbb{F}$. Then there exist constants, c_{ij} for $i \leq j$ and an orthonormal basis, $\{\mathbf{w}_i\}_{i=1}^n$ such that*

$$L = \sum_{j=1}^{n} \sum_{i=1}^{j} c_{ij} \mathbf{w}_i \otimes \mathbf{w}_j$$

The constants, c_{ii} are the eigenvalues of L. Thus the matrix whose ij^{th} entry is c_{ij} is upper triangular.

Proof: If $\dim(H) = 1$, let $H = \text{span}(\mathbf{w})$ where $|\mathbf{w}| = 1$. Then $L\mathbf{w} = k\mathbf{w}$ for some k. Then

$$L = k\mathbf{w} \otimes \mathbf{w}$$

because by definition, $\mathbf{w} \otimes \mathbf{w}(\mathbf{w}) = \mathbf{w}$. Therefore, the theorem holds if H is 1 dimensional.

Now suppose the theorem holds for $n-1 = \dim(H)$. Let $\mathbf{w}_n$ be an eigenvector for L^*. Dividing by its length, it can be assumed $|\mathbf{w}_n| = 1$. Say $L^*\mathbf{w}_n = \mu\mathbf{w}_n$. Using the Gram Schmidt process, there exists an orthonormal basis for H of the form $\{\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}, \mathbf{w}_n\}$. Then

$$(L\mathbf{v}_k, \mathbf{w}_n) = (\mathbf{v}_k, L^*\mathbf{w}_n) = (\mathbf{v}_k, \mu\mathbf{w}_n) = 0,$$

which shows

$$L : H_1 \equiv \text{span}(\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}) \to \text{span}(\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}).$$

Denote by L_1 the restriction of L to H_1. Since H_1 has dimension $n-1$, the induction hypothesis yields an orthonormal basis, $\{\mathbf{w}_1, \cdots, \mathbf{w}_{n-1}\}$ for H_1 such that

$$L_1 = \sum_{j=1}^{n-1} \sum_{i=1}^{j} c_{ij} \mathbf{w}_i \otimes \mathbf{w}_j. \tag{12.3}$$

Then $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ is an orthonormal basis for H because every vector in

$$\text{span}(\mathbf{v}_1, \cdots, \mathbf{v}_{n-1})$$

has the property that its inner product with $\mathbf{w}_n$ is 0 so in particular, this is true for the vectors $\{\mathbf{w}_1, \cdots, \mathbf{w}_{n-1}\}$. Now define c_{in} to be the scalars satisfying

$$L\mathbf{w}_n \equiv \sum_{i=1}^{n} c_{in} \mathbf{w}_i \tag{12.4}$$

and let

$$B \equiv \sum_{j=1}^{n} \sum_{i=1}^{j} c_{ij} \mathbf{w}_i \otimes \mathbf{w}_j.$$

Then by 12.4,

$$B\mathbf{w}_n = \sum_{j=1}^{n} \sum_{i=1}^{j} c_{ij} \mathbf{w}_i \delta_{nj} = \sum_{j=1}^{n} c_{in} \mathbf{w}_i = L\mathbf{w}_n.$$

If $1 \leq k \leq n-1$,

$$B\mathbf{w}_k = \sum_{j=1}^{n} \sum_{i=1}^{j} c_{ij} \mathbf{w}_i \delta_{kj} = \sum_{i=1}^{k} c_{ik} \mathbf{w}_i$$

while from 12.3,

$$L\mathbf{w}_k = L_1\mathbf{w}_k = \sum_{j=1}^{n-1}\sum_{i=1}^{j} c_{ij}\mathbf{w}_i\delta_{jk} = \sum_{i=1}^{k} c_{ik}\mathbf{w}_i.$$

Since $L = B$ on the basis $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$, it follows $L = B$.

It remains to verify the constants, c_{kk} are the eigenvalues of L, solutions of the equation, $\det(\lambda I - L) = 0$. However, the definition of $\det(\lambda I - L)$ is the same as

$$\det(\lambda I - C)$$

where C is the upper triangular matrix which has c_{ij} for $i \leq j$ and zeros elsewhere. This equals 0 if and only if λ is one of the diagonal entries, one of the c_{kk}. ■

Now with the above Schur's theorem, the following diagonalization theorem comes very easily. Recall the following definition.

Definition 12.2.3 *Let $L \in \mathcal{L}(H, H)$ where H is a finite dimensional inner product space. Then L is Hermitian if $L^* = L$.*

Theorem 12.2.4 *Let $L \in \mathcal{L}(H, H)$ where H is an n dimensional inner product space. If L is Hermitian, then all of its eigenvalues λ_k are real and there exists an orthonormal basis of eigenvectors $\{\mathbf{w}_k\}$ such that*

$$L = \sum_k \lambda_k \mathbf{w}_k \otimes \mathbf{w}_k.$$

Proof: By Schur's theorem, Theorem 12.2.2, there exist $l_{ij} \in \mathbb{F}$ such that

$$L = \sum_{j=1}^{n}\sum_{i=1}^{j} l_{ij}\mathbf{w}_i \otimes \mathbf{w}_j$$

Then by Lemma 11.4.2,

$$\begin{aligned}\sum_{j=1}^{n}\sum_{i=1}^{j} l_{ij}\mathbf{w}_i \otimes \mathbf{w}_j &= L = L^* = \sum_{j=1}^{n}\sum_{i=1}^{j} \left(l_{ij}\mathbf{w}_i \otimes \mathbf{w}_j\right)^* \\ &= \sum_{j=1}^{n}\sum_{i=1}^{j} \overline{l_{ij}}\mathbf{w}_j \otimes \mathbf{w}_i = \sum_{i=1}^{n}\sum_{j=1}^{i} \overline{l_{ji}}\mathbf{w}_i \otimes \mathbf{w}_j\end{aligned}$$

By independence, if $i = j$, $l_{ii} = \overline{l_{ii}}$ and so these are all real. If $i < j$, it follows from independence again that $l_{ij} = 0$ because the coefficients corresponding to $i < j$ are all 0 on the right side. Similarly if $i > j$, it follows $l_{ij} = 0$. Letting $\lambda_k = l_{kk}$, this shows

$$L = \sum_k \lambda_k \mathbf{w}_k \otimes \mathbf{w}_k$$

That each of these $\mathbf{w}_k$ is an eigenvector corresponding to λ_k is obvious from the definition of the tensor product. ■

12.3 Spectral Theory Of Self Adjoint Operators

The following theorem is about the eigenvectors and eigenvalues of a self adjoint operator. Such operators are also called Hermitian as in the case of matrices. The proof given generalizes to the situation of a compact self adjoint operator on a Hilbert space and leads to many very useful results. It is also a very elementary proof because it does not use the fundamental theorem of algebra and it contains a way, very important in applications, of finding the eigenvalues. This proof depends more directly on the methods of analysis than the preceding material. Recall the following notation.

Definition 12.3.1 *Let X be an inner product space and let $S \subseteq X$. Then*

$$S^{\perp} \equiv \{x \in X : (x, s) = 0 \text{ for all } s \in S\}.$$

Note that even if S is not a subspace, $S^{\perp}$ is.

Theorem 12.3.2 *Let $A \in \mathcal{L}(X, X)$ be self adjoint (Hermitian) where X is a finite dimensional inner product space of dimension n. Thus $A = A^*$. Then there exists an orthonormal basis of eigenvectors, $\{v_j\}_{j=1}^n$.*

Proof: Consider (Ax, x). This quantity is always a real number because

$$\overline{(Ax, x)} = (x, Ax) = (x, A^*x) = (Ax, x)$$

thanks to the assumption that A is self adjoint. Now define

$$\lambda_1 \equiv \inf\{(Ax, x) : |x| = 1, x \in X_1 \equiv X\}.$$

Claim: λ_1 is finite and there exists $v_1 \in X$ with $|v_1| = 1$ such that $(Av_1, v_1) = \lambda_1$.

Proof of claim: Let $\{u_j\}_{j=1}^n$ be an orthonormal basis for X and for $x \in X$, let $(x_1, \cdots, x_n)$ be defined as the components of the vector x. Thus,

$$x = \sum_{j=1}^{n} x_j u_j.$$

Since this is an orthonormal basis, it follows from the axioms of the inner product that

$$|x|^2 = \sum_{j=1}^{n} |x_j|^2.$$

Thus

$$(Ax, x) = \left(\sum_{k=1}^{n} x_k A u_k, \sum_{j=1} x_j u_j \right) = \sum_{k,j} x_k \overline{x_j} \left(A u_k, u_j \right),$$

a real valued continuous function of $(x_1, \cdots, x_n)$ which is defined on the compact set

$$K \equiv \{(x_1, \cdots, x_n) \in \mathbb{F}^n : \sum_{j=1}^{n} |x_j|^2 = 1\}.$$

Therefore, it achieves its minimum from the extreme value theorem. Then define

$$v_1 \equiv \sum_{j=1}^{n} x_j u_j$$

where $(x_1, \cdots, x_n)$ is the point of K at which the above function achieves its minimum. This proves the claim.

I claim that λ_1 is an eigenvalue and v_1 is an eigenvector. Letting $w \in X_1 \equiv X$, the function of the real variable, t, given by

$$f(t) \equiv \frac{(A(v_1 + tw), v_1 + tw)}{|v_1 + tw|^2} = \frac{(Av_1, v_1) + 2t \operatorname{Re}(Av_1, w) + t^2 (Aw, w)}{|v_1|^2 + 2t \operatorname{Re}(v_1, w) + t^2 |w|^2}$$

achieves its minimum when $t = 0$. Therefore, the derivative of this function evaluated at $t = 0$ must equal zero. Using the quotient rule, this implies, since $|v_1| = 1$ that

$$2 \operatorname{Re}(Av_1, w) |v_1|^2 - 2 \operatorname{Re}(v_1, w)(Av_1, v_1) = 2 (\operatorname{Re}(Av_1, w) - \operatorname{Re}(v_1, w) \lambda_1) = 0.$$

Thus $\operatorname{Re}(Av_1 - \lambda_1 v_1, w) = 0$ for all $w \in X$. This implies $Av_1 = \lambda_1 v_1$. To see this, let $w \in X$ be arbitrary and let θ be a complex number with $|\theta| = 1$ and

$$|(Av_1 - \lambda_1 v_1, w)| = \theta (Av_1 - \lambda_1 v_1, w).$$

Then

$$|(Av_1 - \lambda_1 v_1, w)| = \operatorname{Re}\left(Av_1 - \lambda_1 v_1, \overline{\theta} w\right) = 0.$$

Since this holds for all w, $Av_1 = \lambda_1 v_1$.

Continuing with the proof of the theorem, let $X_2 \equiv \{v_1\}^{\perp}$. This is a closed subspace of X and $A : X_2 \to X_2$ because for $x \in X_2$,

$$(Ax, v_1) = (x, Av_1) = \lambda_1 (x, v_1) = 0.$$

Let

$$\lambda_2 \equiv \inf \{(Ax, x) : |x| = 1, x \in X_2\}$$

As before, there exists $v_2 \in X_2$ such that $Av_2 = \lambda_2 v_2$, $\lambda_1 \leq \lambda_2$. Now let $X_3 \equiv \{v_1, v_2\}^{\perp}$ and continue in this way. As long as $k < n$, it will be the case that $\{v_1, \cdots, v_k\}^{\perp} \neq \{0\}$. This is because for $k < n$ these vectors cannot be a spanning set and so there exists some $w \notin \operatorname{span}(v_1, \cdots, v_k)$. Then letting z be the closest point to w from $\operatorname{span}(v_1, \cdots, v_k)$, it follows that $w - z \in \{v_1, \cdots, v_k\}^{\perp}$. Thus there is an decreasing sequence of eigenvalues $\{\lambda_k\}_{k=1}^{n}$ and a corresponding sequence of eigenvectors, $\{v_1, \cdots, v_n\}$ with this being an orthonormal set. ■

Contained in the proof of this theorem is the following important corollary.

Corollary 12.3.3 *Let $A \in \mathcal{L}(X, X)$ be self adjoint where X is a finite dimensional inner product space. Then all the eigenvalues are real and for $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ the eigenvalues of A, there exists an orthonormal set of vectors $\{u_1, \cdots, u_n\}$ for which*

$$Au_k = \lambda_k u_k.$$

Furthermore,

$$\lambda_k \equiv \inf \{(Ax, x) : |x| = 1, x \in X_k\}$$

where

$$X_k \equiv \{u_1, \cdots, u_{k-1}\}^{\perp}, X_1 \equiv X.$$

Corollary 12.3.4 *Let $A \in \mathcal{L}(X, X)$ be self adjoint (Hermitian) where X is a finite dimensional inner product space. Then the largest eigenvalue of A is given by*

$$\max \{(A\mathbf{x}, \mathbf{x}) : |\mathbf{x}| = 1\} \tag{12.5}$$

and the minimum eigenvalue of A is given by

$$\min \{(A\mathbf{x}, \mathbf{x}) : |\mathbf{x}| = 1\}. \tag{12.6}$$

Proof: The proof of this is just like the proof of Theorem 12.3.2. Simply replace inf with sup and obtain a decreasing list of eigenvalues. This establishes 12.5. The claim 12.6 follows from Theorem 12.3.2. ■

Another important observation is found in the following corollary.

Corollary 12.3.5 *Let $A \in \mathcal{L}(X, X)$ where A is self adjoint. Then $A = \sum_i \lambda_i v_i \otimes v_i$ where $Av_i = \lambda_i v_i$ and $\{v_i\}_{i=1}^n$ is an orthonormal basis.*

Proof : If v_k is one of the orthonormal basis vectors, $Av_k = \lambda_k v_k$. Also,

$$\sum_i \lambda_i v_i \otimes v_i (v_k) = \sum_i \lambda_i v_i (v_k, v_i) = \sum_i \lambda_i \delta_{ik} v_i = \lambda_k v_k.$$

Since the two linear transformations agree on a basis, it follows they must coincide. ■

By Theorem 11.4.5 this says the matrix of A with respect to this basis $\{v_i\}_{i=1}^n$ is the diagonal matrix having the eigenvalues $\lambda_1, \cdots, \lambda_n$ down the main diagonal.

The result of Courant and Fischer which follows resembles Corollary 12.3.3 but is more useful because it does not depend on a knowledge of the eigenvectors.

Theorem 12.3.6 *Let $A \in \mathcal{L}(X, X)$ be self adjoint where X is a finite dimensional inner product space. Then for $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ the eigenvalues of A, there exist orthonormal vectors $\{u_1, \cdots, u_n\}$ for which*

$$Au_k = \lambda_k u_k.$$

Furthermore,

$$\lambda_k \equiv \max_{w_1,\cdots,w_{k-1}} \left\{ \min \left\{ (Ax, x) : |x| = 1, x \in \{w_1, \cdots, w_{k-1}\}^{\perp} \right\} \right\} \tag{12.7}$$

where if $k = 1, \{w_1, \cdots, w_{k-1}\}^{\perp} \equiv X$.

Proof: From Theorem 12.3.2, there exist eigenvalues and eigenvectors with $\{u_1, \cdots, u_n\}$ orthonormal and $\lambda_i \leq \lambda_{i+1}$.

$$(Ax, x) = \sum_{j=1}^n (Ax, u_j) \overline{(x, u_j)} = \sum_{j=1}^n \lambda_j (x, u_j)(u_j, x) = \sum_{j=1}^n \lambda_j |(x, u_j)|^2$$

Recall that $(z, w) = \sum_j (z, u_j) \overline{(w, u_i)}$. Then let $Y = \{w_1, \cdots, w_{k-1}\}^{\perp}$

$$\inf \{(Ax, x) : |x| = 1, x \in Y\} = \inf \left\{ \sum_{j=1}^n \lambda_j |(x, u_j)|^2 : |x| = 1, x \in Y \right\}$$

$$\leq \inf \left\{ \sum_{j=1}^k \lambda_j |(x, u_j)|^2 : |x| = 1, (x, u_j) = 0 \text{ for } j > k, \text{ and } x \in Y \right\}. \tag{12.8}$$

The reason this is so is that the infimum is taken over a smaller set. Therefore, the infimum gets larger. Now 12.8 is no larger than

$$\inf \left\{ \lambda_k \sum_{j=1}^n |(x, u_j)|^2 : |x| = 1, (x, u_j) = 0 \text{ for } j > k, \text{ and } x \in Y \right\} \leq \lambda_k$$

because since $\{u_1, \cdots, u_n\}$ is an orthonormal basis, $|x|^2 = \sum_{j=1}^n |(x, u_j)|^2$. It follows, since $\{w_1, \cdots, w_{k-1}\}$ is arbitrary,

$$\sup_{w_1,\cdots,w_{k-1}} \left\{ \inf \left\{ (Ax, x) : |x| = 1, x \in \{w_1, \cdots, w_{k-1}\}^{\perp} \right\} \right\} \leq \lambda_k. \tag{12.9}$$

Then from Corollary 12.3.3,

$$\lambda_k = \inf \left\{ (Ax, x) : |x| = 1, x \in \{u_1, \cdots, u_{k-1}\}^{\perp} \right\} \leq$$

$$\sup_{w_1,\cdots,w_{k-1}} \left\{ \inf \left\{ (Ax, x) : |x| = 1, x \in \{w_1, \cdots, w_{k-1}\}^{\perp} \right\} \right\} \leq \lambda_k$$

Hence these are all equal and this proves the theorem. ■

The following corollary is immediate.

Corollary 12.3.7 *Let $A \in \mathcal{L}(X, X)$ be self adjoint where X is a finite dimensional inner product space. Then for $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ the eigenvalues of A, there exist orthonormal vectors $\{u_1, \cdots, u_n\}$ for which*

$$Au_k = \lambda_k u_k.$$

Furthermore,

$$\lambda_k \equiv \max_{w_1,\cdots,w_{k-1}} \left\{ \min \left\{ \frac{(Ax,x)}{|x|^2} : x \neq 0, x \in \{w_1,\cdots,w_{k-1}\}^{\perp} \right\} \right\} \tag{12.10}$$

where if $k = 1, \{w_1,\cdots,w_{k-1}\}^{\perp} \equiv X$.

Here is a version of this for which the roles of max and min are reversed.

Corollary 12.3.8 *Let* $A \in \mathcal{L}(X,X)$ *be self adjoint where* X *is a finite dimensional inner product space. Then for* $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ *the eigenvalues of* A, *there exist orthonormal vectors* $\{u_1,\cdots,u_n\}$ *for which*

$$Au_k = \lambda_k u_k.$$

Furthermore,

$$\lambda_k \equiv \min_{w_1,\cdots,w_{n-k}} \left\{ \max \left\{ \frac{(Ax,x)}{|x|^2} : x \neq 0, x \in \{w_1,\cdots,w_{n-k}\}^{\perp} \right\} \right\} \tag{12.11}$$

where if $k = n, \{w_1,\cdots,w_{n-k}\}^{\perp} \equiv X$.

12.4 Positive And Negative Linear Transformations

The notion of a positive definite or negative definite linear transformation is very important in many applications. In particular it is used in versions of the second derivative test for functions of many variables. Here the main interest is the case of a linear transformation which is an $n \times n$ matrix but the theorem is stated and proved using a more general notation because all these issues discussed here have interesting generalizations to functional analysis.

Definition 12.4.1 *A self adjoint* $A \in \mathcal{L}(X,X)$, *is positive definite if whenever* $\mathbf{x} \neq \mathbf{0}$, $(A\mathbf{x},\mathbf{x}) > 0$ *and* A *is negative definite if for all* $\mathbf{x} \neq \mathbf{0}$, $(A\mathbf{x},\mathbf{x}) < 0$. A *is positive semidefinite or just nonnegative for short if for all* $\mathbf{x}$, $(A\mathbf{x},\mathbf{x}) \geq 0$. A *is negative semidefinite or nonpositive for short if for all* $\mathbf{x}$, $(A\mathbf{x},\mathbf{x}) \leq 0$.

The following lemma is of fundamental importance in determining which linear transformations are positive or negative definite.

Lemma 12.4.2 *Let* X *be a finite dimensional inner product space. A self adjoint* $A \in \mathcal{L}(X,X)$ *is positive definite if and only if all its eigenvalues are positive and negative definite if and only if all its eigenvalues are negative. It is positive semidefinite if all the eigenvalues are nonnegative and it is negative semidefinite if all the eigenvalues are nonpositive.*

Proof: Suppose first that A is positive definite and let λ be an eigenvalue. Then for $\mathbf{x}$ an eigenvector corresponding to λ, $\lambda(\mathbf{x},\mathbf{x}) = (\lambda\mathbf{x},\mathbf{x}) = (A\mathbf{x},\mathbf{x}) > 0$. Therefore, $\lambda > 0$ as claimed.

Now suppose all the eigenvalues of A are positive. From Theorem 12.3.2 and Corollary 12.3.5, $A = \sum_{i=1}^n \lambda_i \mathbf{u}_i \otimes \mathbf{u}_i$ where the λ_i are the positive eigenvalues and $\{\mathbf{u}_i\}$ are an orthonormal set of eigenvectors. Therefore, letting $\mathbf{x} \neq \mathbf{0}$,

$$\begin{aligned}(A\mathbf{x},\mathbf{x}) &= \left(\left(\sum_{i=1}^n \lambda_i \mathbf{u}_i \otimes \mathbf{u}_i \right) \mathbf{x}, \mathbf{x} \right) = \left(\sum_{i=1}^n \lambda_i \mathbf{u}_i (\mathbf{x},\mathbf{u}_i), \mathbf{x} \right) \\ &= \left(\sum_{i=1}^n \lambda_i (\mathbf{x},\mathbf{u}_i)(\mathbf{u}_i,\mathbf{x}) \right) = \sum_{i=1}^n \lambda_i |(\mathbf{u}_i,\mathbf{x})|^2 > 0\end{aligned}$$

because, since $\{\mathbf{u}_i\}$ is an orthonormal basis, $|\mathbf{x}|^2 = \sum_{i=1}^n |(\mathbf{u}_i,\mathbf{x})|^2$.

To establish the claim about negative definite, it suffices to note that A is negative definite if and only if $-A$ is positive definite and the eigenvalues of A are (-1) times the eigenvalues of $-A$. The claims about positive semidefinite and negative semidefinite are obtained similarly. ■

The next theorem is about a way to recognize whether a self adjoint $n \times n$ complex matrix A is positive or negative definite without having to find the eigenvalues. In order to state this theorem, here is some notation.

Definition 12.4.3 *Let A be an $n \times n$ matrix. Denote by A_k the $k \times k$ matrix obtained by deleting the $k+1, \cdots, n$ columns and the $k+1, \cdots, n$ rows from A. Thus $A_n = A$ and A_k is the $k \times k$ submatrix of A which occupies the upper left corner of A. The determinants of these submatrices are called the principle minors.*

The following theorem is proved in [8]. For the sake of simplicity, we state this for real matrices since this is also where the main interest lies.

Theorem 12.4.4 *Let A be a self adjoint $n \times n$ matrix. Then A is positive definite if and only if $\det(A_k) > 0$ for every $k = 1, \cdots, n$.*

Proof: This theorem is proved by induction on n. It is clearly true if $n = 1$. Suppose then that it is true for $n-1$ where $n \geq 2$. Since $\det(A) > 0$, it follows that all the eigenvalues are nonzero. Are they all positive? Suppose not. Then there is some even number of them which are negative, even because the product of all the eigenvalues is known to be positive, equaling $\det(A)$. Pick two, λ_1 and λ_2 and let $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ where $\mathbf{u}_i \neq \mathbf{0}$ for $i = 1, 2$ and $(\mathbf{u}_1, \mathbf{u}_2) = 0$. Now if $\mathbf{y} \equiv \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2$ is an element of $\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$, then since these are eigenvalues and $(\mathbf{u}_1, \mathbf{u}_2)_{\mathbb{R}^n} = 0$, a short computation shows

$$(A(\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2), \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2) = |\alpha_1|^2 \lambda_1 |\mathbf{u}_1|^2 + |\alpha_2|^2 \lambda_2 |\mathbf{u}_2|^2 < 0.$$

Now letting $\mathbf{x} \in \mathbb{R}^{n-1}$, $\mathbf{x} \neq \mathbf{0}$, the induction hypothesis implies

$$(\mathbf{x}^T, 0) A \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} = \mathbf{x}^T A_{n-1} \mathbf{x} = (A_{n-1}\mathbf{x}, \mathbf{x}) > 0.$$

The dimension of $\{\mathbf{z} \in \mathbb{R}^n : z_n = 0\}$ is $n-1$ and the dimension of $\text{span}(\mathbf{u}_1, \mathbf{u}_2) = 2$ and so there must be some nonzero $\mathbf{x} \in \mathbb{R}^n$ which is in both of these subspaces of $\mathbb{R}^n$. However, the first computation would require that $(A\mathbf{x}, \mathbf{x}) < 0$ while the second would require that $(A\mathbf{x}, \mathbf{x}) > 0$. This contradiction shows that all the eigenvalues must be positive. This proves the if part of the theorem.

To show the converse, note that, as above, $(A\mathbf{x}, \mathbf{x}) = \mathbf{x}^T A\mathbf{x}$. Suppose that A is positive definite. Then this is equivalent to having

$$\mathbf{x}^T A\mathbf{x} \geq \delta \|\mathbf{x}\|^2$$

Note that for $\mathbf{x} \in \mathbb{R}^k$,

$$\begin{pmatrix} \mathbf{x}^T & \mathbf{0} \end{pmatrix} A \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} = \mathbf{x}^T A_k \mathbf{x} \geq \delta \|\mathbf{x}\|^2$$

From Lemma 12.4.2, this implies that all the eigenvalues of A_k are positive. Hence from Lemma 12.4.2, it follows that $\det(A_k) > 0$, being the product of its eigenvalues. ■

Corollary 12.4.5 *Let A be a self adjoint $n \times n$ matrix. Then A is negative definite if and only if* $\det(A_k)(-1)^k > 0$ *for every $k = 1, \cdots, n$.*

Proof: This is immediate from the above theorem by noting that, as in the proof of Lemma 12.4.2, A is negative definite if and only if $-A$ is positive definite. Therefore, $\det(-A_k) > 0$ for all $k = 1, \cdots, n$, is equivalent to having A negative definite. However, $\det(-A_k) = (-1)^k \det(A_k)$. ■

12.5 The Square Root

With the above theory, it is possible to take fractional powers of certain elements of $\mathcal{L}(X, X)$ where X is a finite dimensional inner product space. I will give two treatments of this, the first pertaining to the square root only and the second more generally pertaining to the k^{th} root of a self adjoint nonnegative matrix.

Theorem 12.5.1 *Let $A \in \mathcal{L}(X, X)$ be self adjoint and nonnegative. Then there exists a unique self adjoint nonnegative $B \in \mathcal{L}(X, X)$ such that $B^2 = A$ and B commutes with every element of $\mathcal{L}(X, X)$ which commutes with A.*

Proof: By Theorem 12.3.2, there exists an orthonormal basis of eigenvectors of A, say $\{v_i\}_{i=1}^n$ such that $Av_i = \lambda_i v_i$. Therefore, by Theorem 12.2.4, $A = \sum_i \lambda_i v_i \otimes v_i$ where each $\lambda_i \geq 0$.

Now by Lemma 12.4.2, each $\lambda_i \geq 0$. Therefore, it makes sense to define

$$B \equiv \sum_i \lambda_i^{1/2} v_i \otimes v_i.$$

It is easy to verify that

$$(v_i \otimes v_i)(v_j \otimes v_j) = \begin{cases} 0 \text{ if } i \neq j \\ v_i \otimes v_i \text{ if } i = j \end{cases}.$$

Therefore, a short computation verifies that $B^2 = \sum_i \lambda_i v_i \otimes v_i = A$. If C commutes with A, then for some c_{ij},

$$C = \sum_{ij} c_{ij} v_i \otimes v_j$$

and so since they commute,

$$\sum_{i,j,k} c_{ij} v_i \otimes v_j \lambda_k v_k \otimes v_k = \sum_{i,j,k} c_{ij} \lambda_k \delta_{jk} v_i \otimes v_k = \sum_{i,k} c_{ik} \lambda_k v_i \otimes v_k$$

$$\begin{aligned} &= \sum_{i,j,k} c_{ij}\lambda_k v_k \otimes v_k v_i \otimes v_j = \sum_{i,j,k} c_{ij}\lambda_k \delta_{ki} v_k \otimes v_j = \sum_{j,k} c_{kj}\lambda_k v_k \otimes v_j \\ &= \sum_{k,i} c_{ik}\lambda_i v_i \otimes v_k \end{aligned}$$

Then by independence,

$$c_{ik}\lambda_i = c_{ik}\lambda_k$$

Therefore, $c_{ik}\lambda_i^{1/2} = c_{ik}\lambda_k^{1/2}$ which amounts to saying that B also commutes with C. It is clear that this operator is self adjoint. This proves existence.

Suppose B_1 is another square root which is self adjoint, nonnegative and commutes with every linear transformation which commutes with A. Since both B, B_1 are nonnegative,

$$\begin{aligned} (B (B - B_1) x, (B - B_1) x) &\geq 0, \\ (B_1 (B - B_1) x, (B - B_1) x) &\geq 0 \end{aligned} \tag{12.12}$$

Now, adding these together, and using the fact that the two commute,

$$\left(\left(B^2 - B_1^2\right) x, (B - B_1) x\right) = ((A - A) x, (B - B_1) x) = 0.$$

It follows that both inner products in 12.12 equal 0. Next use the existence part of this to take the square root of B and B_1 which is denoted by $\sqrt{B}, \sqrt{B_1}$ respectively. Then

$$\begin{aligned} 0 &= \left(\sqrt{B} (B - B_1) x, \sqrt{B} (B - B_1) x\right) \\ 0 &= \left(\sqrt{B_1} (B - B_1) x, \sqrt{B_1} (B - B_1) x\right) \end{aligned}$$

which implies $\sqrt{B} (B - B_1) x = \sqrt{B_1} (B - B_1) x = 0$. Thus also,

$$B (B - B_1) x = B_1 (B - B_1) x = 0$$

Hence

$$0 = (B (B - B_1) x - B_1 (B - B_1) x, x) = ((B - B_1) x, (B - B_1) x)$$

and so, since x is arbitrary, $B_1 = B$. ■

12.6 Fractional Powers

The main result is the following theorem.

Theorem 12.6.1 *Let A be a self adjoint and nonnegative $n \times n$ matrix (all eigenvalues are nonnegative) and let k be a positive integer. Then there exists a unique self adjoint nonnegative matrix B such that $B^k = A$.*

Proof: By Theorem 12.3.2 or Corollary 6.4.12, there exists an orthonormal basis of eigenvectors of A, say $\{v_i\}_{i=1}^n$ such that $Av_i = \lambda_i v_i$ with each λ_i real. In particular, there exists a unitary matrix U such that

$$U^* A U = D, \quad A = U D U^*$$

where D has nonnegative diagonal entries. Define B in the obvious way.

$$B \equiv U D^{1/k} U^*$$

Then it is clear that B is self adjoint and nonnegative. Also it is clear that $B^k = A$. What of uniqueness? Let $p(t)$ be a polynomial whose graph contains the ordered pairs $\left(\lambda_i, \lambda_i^{1/k}\right)$ where the λ_i are the diagonal entries of D, the eigenvalues of A. Then

$$p(A) = U P(D) U^* = U D^{1/k} U^* \equiv B$$

Suppose then that $C^k = A$ and C is also self adjoint and nonnegative.

$$CB = Cp(A) = Cp\left(C^k\right) = p\left(C^k\right)C = p(A)C = BC$$

and so $\{B, C\}$ is a commuting family of non defective matrices. By Theorem 12.1.9 this family of matrices is simultaneously diagonalizable. Hence there exists a single S such that

$$S^{-1}BS = D_B, \quad S^{-1}CS = D_C$$

Where D_C, D_B denote diagonal matrices. Hence, raising to the power k, it follows that

$$A = B^k = SD_B^k S^{-1}, \quad A = C^k = SD_C^k S^{-1}$$

Hence

$$SD_B^k S^{-1} = SD_C^k S^{-1}$$

and so $D_B^k = D_C^k$. Since the entries of the two diagonal matrices are nonnegative, this implies $D_B = D_C$ and so $S^{-1}BS = S^{-1}CS$ which shows $B = C$. ■

A similar result holds for a general finite dimensional inner product space. See Problem 22 in the exercises.

12.7 Square Roots And Polar Decompositions

An application of Theorem 12.3.2, is the following fundamental result, important in geometric measure theory and continuum mechanics. It is sometimes called the right polar decomposition. The notation used is that which is seen in continuum mechanics, see for example Gurtin [12]. Don't confuse the U in this theorem with a unitary transformation. It is not so. When the following theorem is applied in continuum mechanics, F is normally the deformation gradient, the derivative of a nonlinear map from some subset of three dimensional space to three dimensional space. In this context, U is called the right Cauchy Green strain tensor. It is a measure of how a body is stretched independent of rigid motions. First, here is a simple lemma.

Lemma 12.7.1 *Suppose $R \in \mathcal{L}(X, Y)$ where X, Y are inner product spaces and R preserves distances. Then $R^*R = I$.*

Proof: Since R preserves distances, $|R\mathbf{u}| = |\mathbf{u}|$ for every $\mathbf{u}$. Let $\mathbf{u}, \mathbf{v}$ be arbitrary vectors in X and let $\theta \in \mathbb{C}$, $|\theta| = 1$, and $\theta(R^*R\mathbf{u} - \mathbf{u}, \mathbf{v}) = |(R^*R\mathbf{u} - \mathbf{u}, \mathbf{v})|$. Therefore from the axioms of the inner product,

$$\begin{aligned}
|\mathbf{u}|^2 + |\mathbf{v}|^2 + 2\operatorname{Re}\theta(\mathbf{u}, \mathbf{v}) &= |\theta\mathbf{u}|^2 + |\mathbf{v}|^2 + \theta(\mathbf{u}, \mathbf{v}) + \bar{\theta}(\mathbf{v}, \mathbf{u}) \\
&= |\theta\mathbf{u} + \mathbf{v}|^2 = (R(\theta\mathbf{u} + \mathbf{v}), R(\theta\mathbf{u} + \mathbf{v}))
\end{aligned}$$

$$= (R\theta\mathbf{u}, R\theta\mathbf{u}) + (R\mathbf{v}, R\mathbf{v}) + (R\theta\mathbf{u}, R\mathbf{v}) + (R\mathbf{v}, R\theta\mathbf{u})$$

$$\begin{aligned}
&= |\theta\mathbf{u}|^2 + |\mathbf{v}|^2 + \theta(R^*R\mathbf{u}, \mathbf{v}) + \bar{\theta}(\mathbf{v}, R^*R\mathbf{u}) \\
&= |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2\operatorname{Re}\theta(R^*R\mathbf{u}, \mathbf{v})
\end{aligned}$$

and so for all $\mathbf{u}, \mathbf{v}$,

$$2\operatorname{Re}\theta(R^*R\mathbf{u} - \mathbf{u}, \mathbf{v}) = 2|(R^*R\mathbf{u} - \mathbf{u}, \mathbf{v})| = 0$$

Now let $\mathbf{v} = R^*R\mathbf{u} - \mathbf{u}$. It follows that $R^*R\mathbf{u} - \mathbf{u} = \mathbf{0}$. ■

The decomposition in the following is called the right polar decomposition.

Theorem 12.7.2 *Let X be a inner product space of dimension n and let Y be a inner product space of dimension $m \geq n$ and let $F \in \mathcal{L}(X, Y)$. Then there exists $R \in \mathcal{L}(X, Y)$ and $U \in \mathcal{L}(X, X)$ such that*

$$F = RU, \ U = U^*, (U \text{ is Hermitian}),$$

all eigenvalues of U *are non negative,*

$$U^2 = F^*F, R^*R = I,$$

and $|R\mathbf{x}| = |\mathbf{x}|$.

Proof: $(F^*F)^* = F^*F$ and so by Theorem 12.3.2, there is an orthonormal basis of eigenvectors, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ such that

$$F^*F\mathbf{v}_i = \lambda_i \mathbf{v}_i, \ F^*F = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i.$$

It is also clear that $\lambda_i \geq 0$ because

$$\lambda_i (\mathbf{v}_i, \mathbf{v}_i) = (F^*F\mathbf{v}_i, \mathbf{v}_i) = (F\mathbf{v}_i, F\mathbf{v}_i) \geq 0.$$

Let

$$U \equiv \sum_{i=1}^{n} \lambda_i^{1/2} \mathbf{v}_i \otimes \mathbf{v}_i.$$

Then $U^2 = F^*F$, $U = U^*$, and the eigenvalues of U, $\left\{\lambda_i^{1/2}\right\}_{i=1}^{n}$ are all non negative.

Let $\{U\mathbf{x}_1, \cdots, U\mathbf{x}_r\}$ be an orthonormal basis for $U(X)$. By the Gram Schmidt procedure there exists an extension to an orthonormal basis for X,

$$\{U\mathbf{x}_1, \cdots, U\mathbf{x}_r, \mathbf{y}_{r+1}, \cdots, \mathbf{y}_n\}.$$

Next note that $\{F\mathbf{x}_1, \cdots, F\mathbf{x}_r\}$ is also an orthonormal set of vectors in Y because

$$(F\mathbf{x}_k, F\mathbf{x}_j) = (F^*F\mathbf{x}_k, \mathbf{x}_j) = \left(U^2\mathbf{x}_k, \mathbf{x}_j\right) = (U\mathbf{x}_k, U\mathbf{x}_j) = \delta_{jk}.$$

By the Gram Schmidt procedure, there exists an extension of $\{F\mathbf{x}_1, \cdots, F\mathbf{x}_r\}$ to an orthonormal basis for Y,

$$\{F\mathbf{x}_1, \cdots, F\mathbf{x}_r, \mathbf{z}_{r+1}, \cdots, \mathbf{z}_m\}.$$

Since $m \geq n$, there are at least as many $\mathbf{z}_k$ as there are $\mathbf{y}_k$. Now for $\mathbf{x} \in X$, since

$$\{U\mathbf{x}_1, \cdots, U\mathbf{x}_r, \mathbf{y}_{r+1}, \cdots, \mathbf{y}_n\}$$

is an orthonormal basis for X, there exist unique scalars

$$c_1, \cdots, c_r, d_{r+1}, \cdots, d_n$$

such that

$$\mathbf{x} = \sum_{k=1}^{r} c_k U\mathbf{x}_k + \sum_{k=r+1}^{n} d_k \mathbf{y}_k$$

Define

$$R\mathbf{x} \equiv \sum_{k=1}^{r} c_k F\mathbf{x}_k + \sum_{k=r+1}^{n} d_k \mathbf{z}_k \tag{12.13}$$

Thus

$$|R\mathbf{x}|^2 = \sum_{k=1}^{r} |c_k|^2 + \sum_{k=r+1}^{n} |d_k|^2 = |\mathbf{x}|^2.$$

Therefore, by Lemma 12.7.1 $R^*R = I$.

Then also there exist unique scalars b_k such that for a given $\mathbf{x} \in X$,

$$U\mathbf{x} = \sum_{k=1}^{r} b_k U\mathbf{x}_k \tag{12.14}$$

and so from 12.13,

$$RU\mathbf{x} = \sum_{k=1}^{r} b_k F\mathbf{x}_k = F\left(\sum_{k=1}^{r} b_k \mathbf{x}_k\right)$$

Is $F\left(\sum_{k=1}^{r} b_k \mathbf{x}_k\right) = F(\mathbf{x})$?

$$\left(F\left(\sum_{k=1}^{r} b_k \mathbf{x}_k\right) - F(\mathbf{x}), F\left(\sum_{k=1}^{r} b_k \mathbf{x}_k\right) - F(\mathbf{x})\right)$$

$$\begin{aligned}
&= \left((F^*F)\left(\sum_{k=1}^{r} b_k \mathbf{x}_k - \mathbf{x}\right), \left(\sum_{k=1}^{r} b_k \mathbf{x}_k - \mathbf{x}\right)\right) \\
&= \left(U^2\left(\sum_{k=1}^{r} b_k \mathbf{x}_k - \mathbf{x}\right), \left(\sum_{k=1}^{r} b_k \mathbf{x}_k - \mathbf{x}\right)\right) \\
&= \left(U\left(\sum_{k=1}^{r} b_k \mathbf{x}_k - \mathbf{x}\right), U\left(\sum_{k=1}^{r} b_k \mathbf{x}_k - \mathbf{x}\right)\right) \\
&= \left(\sum_{k=1}^{r} b_k U\mathbf{x}_k - U\mathbf{x}, \sum_{k=1}^{r} b_k U\mathbf{x}_k - U\mathbf{x}\right) = 0
\end{aligned}$$

Because from 12.14, $U\mathbf{x} = \sum_{k=1}^{r} b_k U\mathbf{x}_k$. Therefore, $RU\mathbf{x} = F\left(\sum_{k=1}^{r} b_k \mathbf{x}_k\right) = F(\mathbf{x})$. ■

The following corollary follows as a simple consequence of this theorem. It is called the left polar decomposition.

Corollary 12.7.3 *Let $F \in \mathcal{L}(X,Y)$ and suppose $n \geq m$ where X is a inner product space of dimension n and Y is a inner product space of dimension m. Then there exists a Hermitian $U \in \mathcal{L}(X,X)$, and an element of $\mathcal{L}(X,Y)$, R, such that*

$$F = UR, \; RR^* = I.$$

Proof: Recall that $L^{**} = L$ and $(ML)^* = L^*M^*$. Now apply Theorem 12.7.2 to $F^* \in \mathcal{L}(Y,X)$. Thus, $F^* = R^*U$ where R^* and U satisfy the conditions of that theorem. Then $F = UR$ and $RR^* = R^{**}R^* = I$. ■

The following existence theorem for the polar decomposition of an element of $\mathcal{L}(X,X)$ is a corollary.

Corollary 12.7.4 *Let $F \in \mathcal{L}(X,X)$. Then there exists a Hermitian $W \in \mathcal{L}(X,X)$, and a unitary matrix Q such that $F = WQ$, and there exists a Hermitian $U \in \mathcal{L}(X,X)$ and a unitary R, such that $F = RU$.*

This corollary has a fascinating relation to the question whether a given linear transformation is normal. Recall that an $n \times n$ matrix A, is normal if $AA^* = A^*A$. Retain the same definition for an element of $\mathcal{L}(X,X)$.

Theorem 12.7.5 *Let $F \in \mathcal{L}(X,X)$. Then F is normal if and only if in Corollary 12.7.4 $RU = UR$ and $QW = WQ$.*

Proof: I will prove the statement about $RU = UR$ and leave the other part as an exercise. First suppose that $RU = UR$ and show F is normal. To begin with,

$$UR^* = (RU)^* = (UR)^* = R^*U.$$

Therefore,

$$\begin{aligned} F^*F &= UR^*RU = U^2 \\ FF^* &= RUUR^* = URR^*U = U^2 \end{aligned}$$

which shows F is normal.

Now suppose F is normal. Is $RU = UR$? Since F is normal,

$$FF^* = RUUR^* = RU^2R^*$$

and

$$F^*F = UR^*RU = U^2.$$

Therefore, $RU^2R^* = U^2$, and both are nonnegative and self adjoint. Therefore, the square roots of both sides must be equal by the uniqueness part of the theorem on fractional powers. It follows that the square root of the first, RUR^* must equal the square root of the second, U. Therefore, $RUR^* = U$ and so $RU = UR$. This proves the theorem in one case. The other case in which W and Q commute is left as an exercise. ■

12.8 An Application To Statistics

A random vector is a function $\mathbf{X} : \Omega \to \mathbb{R}^p$ where Ω is a probability space. This means that there exists a σ algebra of measurable sets $\mathcal{F}$ and a probability measure $P : \mathcal{F} \to [0,1]$. In practice, people often don't worry too much about the underlying probability space and instead pay more attention to the distribution measure of the random variable. For E a suitable subset of $\mathbb{R}^p$, this measure gives the probability that $\mathbf{X}$ has values in E. There are often excellent reasons for believing that a random vector is normally distributed. This means that the probability that $\mathbf{X}$ has values in a set E is given by

$$\int_E \frac{1}{(2\pi)^{p/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m})^* \Sigma^{-1} (\mathbf{x}-\mathbf{m})\right) d\mathbf{x}$$

The expression in the integral is called the normal probability density function. There are two parameters, $\mathbf{m}$ and Σ where $\mathbf{m}$ is called the mean and Σ is called the covariance matrix. It is a symmetric matrix which has all real eigenvalues which are all positive. While it may be reasonable to assume this is the distribution, in general, you won't know $\mathbf{m}$ and Σ and in order to use this formula to predict anything, you would need to know these quantities. I am following a nice discussion given in Wikipedia which makes use of the existence of square roots.

What people do to estimate these is to take n independent observations $\mathbf{x}_1, \cdots, \mathbf{x}_n$ and try to predict what $\mathbf{m}$ and Σ should be based on these observations. One criterion used for making this determination is the method of maximum likelihood. In this method, you seek to choose the two parameters in such a way as to maximize the likelihood which is given as

$$\prod_{i=1}^{n} \frac{1}{\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1} (\mathbf{x}_i - \mathbf{m})\right).$$

For convenience the term $(2\pi)^{p/2}$ was ignored. Maximizing the above is equivalent to maximizing the ln of the above. So taking ln,

$$\frac{n}{2} \ln\left(\det\left(\Sigma^{-1}\right)\right) - \frac{1}{2} \sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1} (\mathbf{x}_i - \mathbf{m})$$

Note that the above is a function of the entries of $\mathbf{m}$. Take the partial derivative with respect to m_l. Since the matrix Σ^{-1} is symmetric this implies

$$\sum_{i=1}^{n} \sum_{r} (x_{ir} - m_r) \Sigma_{rl}^{-1} = 0 \text{ each } l.$$

Written in terms of vectors,

$$\sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1} = \mathbf{0}$$

and so, multiplying by Σ on the right and then taking adjoints, this yields

$$\sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{m}) = \mathbf{0},\ n\mathbf{m} = \sum_{i=1}^{n} \mathbf{x}_i,\ \ \mathbf{m} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i \equiv \bar{\mathbf{x}}.$$

Now that $\mathbf{m}$ is determined, it remains to find the best estimate for Σ. $(\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1} (\mathbf{x}_i - \mathbf{m})$ is a scalar, so since $\operatorname{trace}(AB) = \operatorname{trace}(BA)$,

$$\begin{aligned} (\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1} (\mathbf{x}_i - \mathbf{m}) &= \operatorname{trace}\left((\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1} (\mathbf{x}_i - \mathbf{m})\right) \\ &= \operatorname{trace}\left((\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1}\right) \end{aligned}$$

Therefore, the thing to maximize is

$$\begin{aligned} & n \ln\left(\det\left(\Sigma^{-1}\right)\right) - \sum_{i=1}^{n} \operatorname{trace}\left((\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^* \Sigma^{-1}\right) \\ = \ & n \ln\left(\det\left(\Sigma^{-1}\right)\right) - \operatorname{trace}\left(\overbrace{\left(\sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^*\right)}^{S} \Sigma^{-1}\right) \end{aligned}$$

We assume that S has rank p. Thus it is a self adjoint matrix which has all positive eigenvalues. Therefore, from the property of the trace, the thing to maximize is

$$n \ln \left(\det \left(\Sigma^{-1}\right)\right) - \text{trace}\left(S^{1/2}\Sigma^{-1}S^{1/2}\right)$$

Now let $B = S^{1/2}\Sigma^{-1}S^{1/2}$. Then B is positive and self adjoint also and so there exists U unitary such that $B = U^*DU$ where D is the diagonal matrix having the positive scalars $\lambda_1, \cdots, \lambda_p$ down the main diagonal. Solving for Σ^{-1} in terms of B, this yields $S^{-1/2}BS^{-1/2} = \Sigma^{-1}$ and so

$$\ln \left(\det \left(\Sigma^{-1}\right)\right) = \ln \left(\det \left(S^{-1/2}\right) \det (B) \det \left(S^{-1/2}\right)\right) = \ln \left(\det \left(S^{-1}\right)\right) + \ln (\det (B))$$

which yields

$$C(S) + n \ln (\det (B)) - \text{trace}(B)$$

as the thing to maximize. Of course this yields

$$\begin{aligned} & C(S) + n \ln \left(\prod_{i=1}^{p} \lambda_i\right) - \sum_{i=1}^{p} \lambda_i \\ = \quad & C(S) + n \sum_{i=1}^{p} \ln (\lambda_i) - \sum_{i=1}^{p} \lambda_i \end{aligned}$$

as the quantity to be maximized. To do this, take $\partial/\partial\lambda_k$ and set equal to 0. This yields $\lambda_k = n$. Therefore, from the above, $B = U^*nIU = nI$. Also from the above,

$$B^{-1} = \frac{1}{n}I = S^{-1/2}\Sigma S^{-1/2}$$

and so

$$\Sigma = \frac{1}{n}S = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^*$$

This has shown that the maximum likelihood estimates are

$$\mathbf{m} = \bar{\mathbf{x}} \equiv \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i,\ \Sigma = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^*.$$

12.9 The Singular Value Decomposition

In this section, A will be an $m \times n$ matrix. To begin with, here is a simple lemma.

Lemma 12.9.1 *Let A be an $m \times n$ matrix. Then A^*A is self adjoint and all its eigenvalues are nonnegative.*

Proof: It is obvious that A^*A is self adjoint. Suppose $A^*A\mathbf{x} = \lambda\mathbf{x}$. Then $\lambda|\mathbf{x}|^2 = (\lambda\mathbf{x}, \mathbf{x}) = (A^*A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) \geq 0$. ■

Definition 12.9.2 *Let A be an $m \times n$ matrix. The singular values of A are the square roots of the positive eigenvalues of A^*A.*

With this definition and lemma here is the main theorem on the singular value decomposition. In all that follows, I will write the following partitioned matrix

$$\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$$

where σ denotes an $r \times r$ diagonal matrix of the form

$$\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{pmatrix}$$

and the bottom row of zero matrices in the partitioned matrix, as well as the right columns of zero matrices are each of the right size so that the resulting matrix is $m \times n$. Either could vanish completely. However, I will write it in the above form. It is easy to make the necessary adjustments in the other two cases.

Theorem 12.9.3 *Let A be an $m \times n$ matrix. Then there exist unitary matrices, U and V of the appropriate size such that*

$$U^*AV = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$$

where σ is of the form

$$\sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{pmatrix}$$

for the σ_i the singular values of A, arranged in order of decreasing size.

Proof: By the above lemma and Theorem 12.3.2 there exists an orthonormal basis, $\{\mathbf{v}_i\}_{i=1}^n$ for $\mathbb{F}^n$ such that $A^*A\mathbf{v}_i = \sigma_i^2\mathbf{v}_i$ where $\sigma_i^2 > 0$ for $i = 1, \cdots, k, (\sigma_i > 0)$, and equals zero if $i > k$. Let the eigenvalues σ_i^2 be arranged in decreasing order. It is desired to have

$$AV = U\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$$

and so if $U = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_m \end{pmatrix}$, one needs to have for $j \leq k$, $\sigma_j \mathbf{u}_j = A\mathbf{v}_j$. Thus let

$$\mathbf{u}_j \equiv \sigma_j^{-1} A\mathbf{v}_j, \ j \leq k$$

Then for $i, j \leq k$,

$$\begin{aligned}(\mathbf{u}_i, \mathbf{u}_j) &= \sigma_j^{-1}\sigma_i^{-1}(A\mathbf{v}_i, A\mathbf{v}_j) = \sigma_j^{-1}\sigma_i^{-1}(A^*A\mathbf{v}_i, \mathbf{v}_j) \\ &= \sigma_j^{-1}\sigma_i^{-1}\sigma_i^2(\mathbf{v}_i, \mathbf{v}_j) = \delta_{ij}\end{aligned}$$

Now extend to an orthonormal basis of $\mathbb{F}^m$, $\{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}, \cdots, \mathbf{u}_m\}$. If $i > k$,

$$(A\mathbf{v}_i, A\mathbf{v}_i) = (A^*A\mathbf{v}_i, \mathbf{v}_i) = 0\,(\mathbf{v}_i, \mathbf{v}_i) = 0$$

so $A\mathbf{v}_i = \mathbf{0}$. Then for σ given as above in the statement of the theorem, it follows that

$$AV = U\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}, \ U^*AV = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \ \blacksquare$$

The singular value decomposition has as an immediate corollary the following interesting result.

Corollary 12.9.4 *Let A be an $m \times n$ matrix. Then the rank of A and A^* equals the number of singular values.*

Proof: Since V and U are unitary, they are each one to one and onto and so it follows that

$$\text{rank}(A) = \text{rank}(U^*AV) = \text{rank}\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} = \text{number of singular values.}$$

Also since U, V are unitary,

$$\text{rank}(A^*) = \text{rank}(V^*A^*U) = \text{rank}\left((U^*AV)^*\right)$$

$$= \text{rank}\left(\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}^*\right) = \text{number of singular values.} \ \blacksquare$$

12.10 Approximation In The Frobenius Norm

The Frobenius norm is one of many norms for a matrix. It is arguably the most obvious of all norms. Here is its definition.

Definition 12.10.1 *Let A be a complex $m \times n$ matrix. Then*

$$||A||_F \equiv (\text{trace}(AA^*))^{1/2}$$

Also this norm comes from the inner product

$$(A, B)_F \equiv \text{trace}(AB^*)$$

Thus $||A||_F^2$ is easily seen to equal $\sum_{ij} |a_{ij}|^2$ so essentially, it treats the matrix as a vector in $\mathbb{F}^{m \times n}$.

Lemma 12.10.2 *Let A be an $m \times n$ complex matrix with singular matrix*

$$\Sigma = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$$

with σ as defined above, $U^*AV = \Sigma$. Then

$$||\Sigma||_F^2 = ||A||_F^2 \tag{12.15}$$

and the following hold for the Frobenius norm. If U, V are unitary and of the right size,

$$||UA||_F = ||A||_F, \ ||UAV||_F = ||A||_F. \tag{12.16}$$

Proof: From the definition and letting U, V be unitary and of the right size,

$$||UA||_F^2 \equiv \text{trace}\,(UAA^*U^*) = \text{trace}\,(U^*UAA^*) = \text{trace}\,(AA^*) = ||A||_F^2$$

Also,

$$||AV||_F^2 \equiv \text{trace}\,(AVV^*A^*) = \text{trace}\,(AA^*) = ||A||_F^2.$$

It follows

$$||\Sigma||_F^2 = ||U^*AV||_F^2 = ||AV||_F^2 = ||A||_F^2. \blacksquare$$

Of course, this shows that

$$||A||_F^2 = \sum_i \sigma_i^2,$$

the sum of the squares of the singular values of A.

Why is the singular value decomposition important? It implies

$$A = U \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} V^*$$

where σ is the diagonal matrix having the singular values down the diagonal. Now sometimes A is a huge matrix, 1000×2000 or something like that. This happens in applications to situations where the entries of A describe a picture. What also happens is that most of the singular values are very small. What if you deleted those which were very small, say for all $i \geq l$ and got a new matrix

$$A' \equiv U \begin{pmatrix} \sigma' & 0 \\ 0 & 0 \end{pmatrix} V^*?$$

Then the entries of A' would end up being close to the entries of A but there is much less information to keep track of. This turns out to be very useful. More precisely, letting

$$\sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix},\ U^*AV = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix},$$

$$||A - A'||_F^2 = \left\| U \begin{pmatrix} \sigma - \sigma' & 0 \\ 0 & 0 \end{pmatrix} V^* \right\|_F^2 = \sum_{k=l+1}^{r} \sigma_k^2$$

Thus A is approximated by A' where A' has rank $l < r$. In fact, it is also true that out of all matrices of rank l, this A' is the one which is closest to A in the Frobenius norm. Thus A is approximated by A' where A' has rank $l < r$. In fact, it is also true that out of all matrices of rank l, this A' is the one which is closest to A in the Frobenius norm.

Here is roughly why this is so. Suppose $\tilde{B}$ approximates $A = \begin{pmatrix} \sigma_{r\times r} & 0 \\ 0 & 0 \end{pmatrix}$ as well as possible out of all matrices $\tilde{B}$ having rank no more than $l < r$ the size of the matrix $\sigma_{r\times r}$. Suppose the rank of $\tilde{B}$ is l. Then obviously no column $\mathbf{x}_j$ of $\tilde{B}$ in a basis for the column space can have $j > r$ since if so, the approximation of A could be improved by simply making this column into a zero column. Therefore there are $\begin{pmatrix} r \\ l \end{pmatrix}$ choices for columns for a basis for the column space of $\tilde{B}$. Suppose you pick the first l for instance. Thus the first column of $\tilde{B}$ should be $\sigma_1\mathbf{e}_1$ to make the approximation up to the first column as good as possible. Now consider approximating as well as possible up to the first two columns. Clearly the second column should be $\sigma_2\mathbf{e}_2$ and in this way, the approximation up to the first two columns is exact. Continue this way till the l^{th} column. Then since $\tilde{B}$ has rank l, all other columns should be zero columns since you cannot have a nonzero entry in any diagonal position and keep the rank of $\tilde{B}$ only l. Then since it is desired to get the best approximation of A you wouldn't want any off diagonal nonzero terms either. The square of the error in doing this, picking the first l columns as a basis would be $\sum_{j=l+1}^{r} \sigma_j^2$. On the other hand, if you picked other columns than the first l in the basis for the column space of $\tilde{B}$, you would have a larger error because you would include sums involving the larger singular values. Thus letting σ' denote the $l \times l$ upper left corner of σ, $\tilde{B}$ should be of the form $\begin{pmatrix} \sigma' & 0 \\ 0 & 0 \end{pmatrix}$. For example,

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

is best approximated by the rank 2 matrix

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Now suppose A is an $m \times n$ matrix. Let U, V be unitary and of the right size such that

$$U^*AV = \begin{pmatrix} \sigma_{r\times r} & 0 \\ 0 & 0 \end{pmatrix}$$

Then suppose B approximates A as well as possible in the Frobenius norm. Then you would want

$$\|A - B\| = \|U^*AV - U^*BV\| = \left\| \begin{pmatrix} \sigma_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} - U^*BV \right\|$$

to be as small as possible. Therefore, from the above discussion, you should have to be as small as possible. Therefore, from the above discussion, you should have

$$U^*BV = \begin{pmatrix} \sigma' & 0 \\ 0 & 0 \end{pmatrix}, B = U \begin{pmatrix} \sigma' & 0 \\ 0 & 0 \end{pmatrix} V^*$$

whereas

$$A = U \begin{pmatrix} \sigma_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} V^*$$

12.11 Least Squares And Singular Value Decomposition

The singular value decomposition also has a very interesting connection to the problem of least squares solutions. Recall that it was desired to find $\mathbf{x}$ such that $|A\mathbf{x} - \mathbf{y}|$ is as small as possible. Lemma 11.5.1 shows that there is a solution to this problem which can be found by solving the system $A^*A\mathbf{x} = A^*\mathbf{y}$. Each $\mathbf{x}$ which solves this system solves the minimization problem as was shown in the lemma just mentioned. Now consider this equation for the solutions of the minimization problem in terms of the singular value decomposition.

$$\overbrace{V \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} U^*}^{A^*} \overbrace{U \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} V^*}^{A} \mathbf{x} = \overbrace{V \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} U^*}^{A^*} \mathbf{y}.$$

Therefore, this yields the following upon using block multiplication and multiplying on the left by V^*.

$$\begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix} V^*\mathbf{x} = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} U^*\mathbf{y}. \tag{12.17}$$

One solution to this equation which is very easy to spot is

$$\mathbf{x} = V \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*\mathbf{y}. \tag{12.18}$$

12.12 The Moore Penrose Inverse

The particular solution of the least squares problem given in 12.18 is important enough that it motivates the following definition.

Definition 12.12.1 *Let A be an $m \times n$ matrix. Then the Moore Penrose inverse of A, denoted by A^+ is defined as*

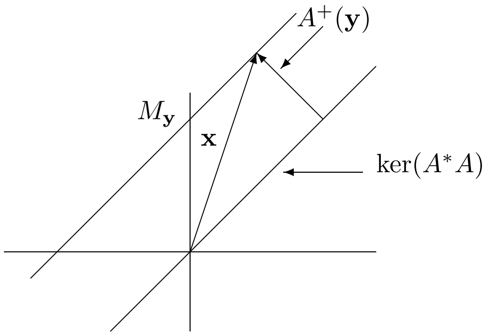
$$A^+ \equiv V \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*.$$

Here

$$U^*AV = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$$

as above.

Thus $A^+\mathbf{y}$ is a solution to the minimization problem to find $\mathbf{x}$ which minimizes $|A\mathbf{x} - \mathbf{y}|$. In fact, one can say more about this. In the following picture $M_{\mathbf{y}}$ denotes the set of least squares solutions $\mathbf{x}$ such that $A^*A\mathbf{x} = A^*\mathbf{y}$.



Then $A^+(\mathbf{y})$ is as given in the picture.

Proposition 12.12.2 *$A^+\mathbf{y}$ is the solution to the problem of minimizing $|A\mathbf{x}-\mathbf{y}|$ for all $\mathbf{x}$ which has smallest norm. Thus*

$$\left|AA^+\mathbf{y}-\mathbf{y}\right| \leq |A\mathbf{x}-\mathbf{y}| \text{ for all } \mathbf{x}$$

and if $\mathbf{x}_1$ satisfies $|A\mathbf{x}_1-\mathbf{y}| \leq |A\mathbf{x}-\mathbf{y}|$ for all $\mathbf{x}$, then $|A^+\mathbf{y}| \leq |\mathbf{x}_1|$.

Proof: Consider $\mathbf{x}$ satisfying 12.17, equivalently $A^*A\mathbf{x} = A^*\mathbf{y}$,

$$\begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix} V^*\mathbf{x} = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} U^*\mathbf{y}$$

which has smallest norm. This is equivalent to making $|V^*\mathbf{x}|$ as small as possible because V^* is unitary and so it preserves norms. For $\mathbf{z}$ a vector, denote by $(\mathbf{z})_k$ the vector in $\mathbb{F}^k$ which consists of the first k entries of $\mathbf{z}$. Then if $\mathbf{x}$ is a solution to 12.17

$$\begin{pmatrix} \sigma^2 (V^*\mathbf{x})_k \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \sigma (U^*\mathbf{y})_k \\ \mathbf{0} \end{pmatrix}$$

and so $(V^*\mathbf{x})_k = \sigma^{-1} (U^*\mathbf{y})_k$. Thus the first k entries of $V^*\mathbf{x}$ are determined. In order to make $|V^*\mathbf{x}|$ as small as possible, the remaining $n-k$ entries should equal zero. Therefore,

$$V^*\mathbf{x} = \begin{pmatrix} (V^*\mathbf{x})_k \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma^{-1} (U^*\mathbf{y})_k \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*\mathbf{y}$$

and so

$$\mathbf{x} = V \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*\mathbf{y} \equiv A^+\mathbf{y} \ \blacksquare$$

Lemma 12.12.3 *The matrix A^+ satisfies the following conditions.*

$$AA^+A = A,\ A^+AA^+ = A^+,\ A^+A \text{ and } AA^+ \text{ are Hermitian.} \tag{12.19}$$

Proof: This is routine. Recall

$$A = U \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} V^*$$

and

$$A^+ = V \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*$$

so you just plug in and verify it works. ■

A much more interesting observation is that A^+ is characterized as being the unique matrix which satisfies 12.19. This is the content of the following Theorem. The conditions are sometimes called the Penrose conditions.

Theorem 12.12.4 *Let A be an $m \times n$ matrix. Then a matrix A_0, is the Moore Penrose inverse of A if and only if A_0 satisfies*

$$AA_0A = A,\ A_0AA_0 = A_0,\ A_0A \text{ and } AA_0 \text{ are Hermitian.} \tag{12.20}$$

Proof: From the above lemma, the Moore Penrose inverse satisfies 12.20. Suppose then that A_0 satisfies 12.20. It is necessary to verify that $A_0 = A^+$. Recall that from the singular value decomposition, there exist unitary matrices, U and V such that

$$U^*AV = \Sigma \equiv \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix},\ A = U\Sigma V^*.$$

Recall that

$$A^+ = V \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*$$

Let

$$A_0 = V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^* \tag{12.21}$$

where P is $r \times r$, the same size as the diagonal matrix composed of the singular values on the main diagonal.

Next use the first equation of 12.20 to write

$$\overbrace{U\Sigma V^*}^{A} \overbrace{V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*}^{A_0} \overbrace{U\Sigma V^*}^{A} = \overbrace{U\Sigma V^*}^{A}.$$

Then multiplying both sides on the left by U^* and on the right by V,

$$\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sigma P \sigma & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \tag{12.22}$$

Therefore, $P = \sigma^{-1}$. From the requirement that AA_0 is Hermitian,

$$\overbrace{U\Sigma V^*}^{A} \overbrace{V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*}^{A_0} = U \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*$$

must be Hermitian. Therefore, it is necessary that

$$\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} \sigma P & \sigma Q \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & \sigma Q \\ 0 & 0 \end{pmatrix}$$

is Hermitian. Then

$$\begin{pmatrix} I & \sigma Q \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ Q^*\sigma & 0 \end{pmatrix}$$

and so $Q = 0$.

Next,

$$\overbrace{V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*}^{A_0} \overbrace{U\Sigma V^*}^{A} = V \begin{pmatrix} P\sigma & 0 \\ R\sigma & 0 \end{pmatrix} V^* = V \begin{pmatrix} I & 0 \\ R\sigma & 0 \end{pmatrix} V^*$$

is Hermitian. Therefore, also

$$\begin{pmatrix} I & 0 \\ R\sigma & 0 \end{pmatrix}$$

is Hermitian. Thus $R = 0$ because

$$\begin{pmatrix} I & 0 \\ R\sigma & 0 \end{pmatrix}^* = \begin{pmatrix} I & \sigma^* R^* \\ 0 & 0 \end{pmatrix}$$

which requires $R\sigma = 0$. Now multiply on right by σ^{-1} to find that $R = 0$.

Use 12.21 and the second equation of 12.20 to write

$$\overbrace{V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*}^{A_0} \overbrace{U\Sigma V^*}^{A} \overbrace{V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*}^{A_0} = \overbrace{V \begin{pmatrix} P & Q \\ R & S \end{pmatrix} U^*}^{A_0}.$$

which implies

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}.$$

This yields from the above in which is was shown that R, Q are both 0

$$\begin{pmatrix} \sigma^{-1} & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & S \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} \tag{12.23}$$

$$= \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & S \end{pmatrix}. \tag{12.24}$$

Therefore, $S = 0$ also and so

$$V^*A_0U \equiv \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

which says

$$A_0 = V \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^* \equiv A^+. \blacksquare$$

The theorem is significant because there is no mention of eigenvalues or eigenvectors in the characterization of the Moore Penrose inverse given in 12.20. It also shows immediately that the Moore Penrose inverse is a generalization of the usual inverse. See Problem 3.

12.13 Exercises

1. Show $(A^*)^* = A$ and $(AB)^* = B^*A^*$.
2. Prove Corollary 12.3.8.
3. Show that if A is an $n \times n$ matrix which has an inverse then $A^+ = A^{-1}$.
4. Using the singular value decomposition, show that for any square matrix A, it follows that A^*A is unitarily similar to AA^*.
5. Let A, B be a $m \times n$ matrices. Define an inner product on the set of $m \times n$ matrices by
$$(A, B)_F \equiv \text{trace}\,(AB^*)\,.$$
Show this is an inner product satisfying all the inner product axioms. Recall for M an $n \times n$ matrix, $\text{trace}\,(M) \equiv \sum_{i=1}^n M_{ii}$. The resulting norm, $||\cdot||_F$ is called the Frobenius norm and it can be used to measure the distance between two matrices.

6. Let A be an $m \times n$ matrix. Show $||A||_F^2 \equiv (A, A)_F = \sum_j \sigma_j^2$ where the σ_j are the singular values of A.

7. If A is a general $n \times n$ matrix having possibly repeated eigenvalues, show there is a sequence $\{A_k\}$ of $n \times n$ matrices having distinct eigenvalues which has the property that the ij^{th} entry of A_k converges to the ij^{th} entry of A for all ij. **Hint:** Use Schur's theorem.

8. Prove the Cayley Hamilton theorem as follows. First suppose A has a basis of eigenvectors $\{\mathbf{v}_k\}_{k=1}^n, A\mathbf{v}_k = \lambda_k \mathbf{v}_k$. Let $p(\lambda)$ be the characteristic polynomial. Show $p(A)\mathbf{v}_k = p(\lambda_k)\mathbf{v}_k = \mathbf{0}$. Then since $\{\mathbf{v}_k\}$ is a basis, it follows $p(A)\mathbf{x} = \mathbf{0}$ for all $\mathbf{x}$ and so $p(A) = 0$. Next in the general case, use Problem 7 to obtain a sequence $\{A_k\}$ of matrices whose entries converge to the entries of A such that A_k has n distinct eigenvalues and therefore by Theorem 6.1.7 A_k has a basis of eigenvectors. Therefore, from the first part and for $p_k(\lambda)$ the characteristic polynomial for A_k, it follows $p_k(A_k) = 0$. Now explain why and the sense in which $\lim_{k\to\infty} p_k(A_k) = p(A)$.

9. Prove that Theorem 12.4.4 and Corollary 12.4.5 can be strengthened so that the condition on the A_k is necessary as well as sufficient. **Hint:** Consider vectors of the form $\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}$ where $\mathbf{x} \in \mathbb{F}^k$.

10. Show directly that if A is an $n \times n$ matrix and $A = A^*$ (A is Hermitian) then all the eigenvalues are real and eigenvectors can be assumed to be real and that eigenvectors associated with distinct eigenvalues are orthogonal, (their inner product is zero).

11. Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be an orthonormal basis for $\mathbb{F}^n$. Let Q be a matrix whose i^{th} column is $\mathbf{v}_i$. Show
$$Q^*Q = QQ^* = I.$$

12. Show that an $n \times n$ matrix Q is unitary if and only if it preserves distances. This means $|Q\mathbf{v}| = |\mathbf{v}|$. This was done in the text but you should try to do it for yourself.

13. Suppose $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ are two orthonormal bases for $\mathbb{F}^n$ and suppose Q is an $n \times n$ matrix satisfying $Q\mathbf{v}_i = \mathbf{w}_i$. Then show Q is unitary. If $|\mathbf{v}| = 1$, show there is a unitary transformation which maps $\mathbf{v}$ to $\mathbf{e}_1$.

14. Finish the proof of Theorem 12.7.5.

15. Let A be a Hermitian matrix so $A = A^*$ and suppose all eigenvalues of A are larger than δ^2. Show
$$(A\mathbf{v}, \mathbf{v}) \geq \delta^2 |\mathbf{v}|^2$$
Where here, the inner product is $(\mathbf{v}, \mathbf{u}) \equiv \sum_{j=1}^n v_j \overline{u_j}$.

16. The discrete Fourier transform maps $\mathbb{C}^n \to \mathbb{C}^n$ as follows.
$$F(\mathbf{x}) = \mathbf{z} \text{ where } z_k = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} e^{-i\frac{2\pi}{n}jk} x_j.$$
Show that F^{-1} exists and is given by the formula
$$F^{-1}(\mathbf{z}) = \mathbf{x} \text{ where } x_j = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} e^{i\frac{2\pi}{n}jk} z_k$$

Here is one way to approach this problem. Note $\mathbf{z} = U\mathbf{x}$ where

$$U = \frac{1}{\sqrt{n}} \begin{pmatrix} e^{-i\frac{2\pi}{n}0\cdot 0} & e^{-i\frac{2\pi}{n}1\cdot 0} & e^{-i\frac{2\pi}{n}2\cdot 0} & \cdots & e^{-i\frac{2\pi}{n}(n-1)\cdot 0} \\ e^{-i\frac{2\pi}{n}0\cdot 1} & e^{-i\frac{2\pi}{n}1\cdot 1} & e^{-i\frac{2\pi}{n}2\cdot 1} & \cdots & e^{-i\frac{2\pi}{n}(n-1)\cdot 1} \\ e^{-i\frac{2\pi}{n}0\cdot 2} & e^{-i\frac{2\pi}{n}1\cdot 2} & e^{-i\frac{2\pi}{n}2\cdot 2} & \cdots & e^{-i\frac{2\pi}{n}(n-1)\cdot 2} \\ \vdots & \vdots & \vdots & & \vdots \\ e^{-i\frac{2\pi}{n}0\cdot(n-1)} & e^{-i\frac{2\pi}{n}1\cdot(n-1)} & e^{-i\frac{2\pi}{n}2\cdot(n-1)} & \cdots & e^{-i\frac{2\pi}{n}(n-1)\cdot(n-1)} \end{pmatrix}$$

Now argue U is unitary and use this to establish the result. To show this verify each row has length 1 and the inner product of two different rows gives 0. Now $U_{kj} = e^{-i\frac{2\pi}{n}jk}$ and so $(U^*)_{kj} = e^{i\frac{2\pi}{n}jk}$.

17. Let f be a periodic function having period 2π. The Fourier series of f is an expression of the form

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx} \equiv \lim_{n\to\infty} \sum_{k=-n}^{n} c_k e^{ikx}$$

and the idea is to find c_k such that the above sequence converges in some way to f. If

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

and you formally multiply both sides by e^{-imx} and then integrate from 0 to 2π, interchanging the integral with the sum without any concern for whether this makes sense, show it is reasonable from this to expect

$$c_m = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{-imx} dx.$$

Now suppose you only know $f(x)$ at equally spaced points $2\pi j/n$ for $j = 0, 1, \cdots, n$. Consider the Riemann sum for this integral obtained from using the left endpoint of the subintervals determined from the partition $\left\{\frac{2\pi}{n}j\right\}_{j=0}^{n}$. How does this compare with the discrete Fourier transform? What happens as $n \to \infty$ to this approximation?

18. Suppose A is a real 3×3 orthogonal matrix (Recall this means $AA^T = A^TA = I$.) having determinant 1. Show it must have an eigenvalue equal to 1. Note this shows there exists a vector $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{x}$. **Hint:** Show first or recall that any orthogonal matrix must preserve lengths. That is, $|A\mathbf{x}| = |\mathbf{x}|$.

19. Let A be a complex $m \times n$ matrix. Using the description of the Moore Penrose inverse in terms of the singular value decomposition, show that

$$\lim_{\delta\to 0+} (A^*A + \delta I)^{-1} A^* = A^+$$

where the convergence happens in the Frobenius norm. Also verify, using the singular value decomposition, that the inverse exists in the above formula. Observe that this shows that the Moore Penrose inverse is unique.

20. Show that $A^+ = (A^*A)^+ A^*$. **Hint:** You might use the description of A^+ in terms of the singular value decomposition.

21. In Theorem 12.6.1. Show that every matrix which commutes with A also commutes with $A^{1/k}$ the unique nonnegative self adjoint k^{th} root.

22. Let X be a finite dimensional inner product space and let $\beta = \{u_1, \cdots, u_n\}$ be an orthonormal basis for X. Let $A \in \mathcal{L}(X, X)$ be self adjoint and nonnegative and let M be its matrix with respect to the given orthonormal basis. Show that M is nonnegative, self adjoint also. Use this to show that A has a unique nonnegative self adjoint k^{th} root.

23. Let A be a complex $m \times n$ matrix having singular value decomposition $U^*AV = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$ as explained above, where σ is $k \times k$. Show that

$$\ker(A) = \text{span}\left(V\mathbf{e}_{k+1}, \cdots, V\mathbf{e}_n\right),$$

the last $n-k$ columns of V.

24. The principal submatrices of an $n \times n$ matrix A are A_k where A_k consists those entries which are in the first k rows and first k columns of A. Suppose A is a real symmetric matrix and that $\mathbf{x} \to \langle A\mathbf{x}, \mathbf{x} \rangle$ is positive definite. This means that if $\mathbf{x} \neq \mathbf{0}$, then $\langle A\mathbf{x}, \mathbf{x} \rangle > 0$. Show that each of the principal submatrices are positive definite. **Hint:** Consider $\begin{pmatrix} \mathbf{x}^T & \mathbf{0} \end{pmatrix} A \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}$ where $\mathbf{x}$ consists of k entries.

25. ↑Show that if A is a symmetric positive definite $n \times n$ real matrix, then A has an LU factorization with the property that each entry on the main diagonal in U is positive. **Hint:** This is pretty clear if A is 1×1. Assume true for $(n-1) \times (n-1)$. Then

$$A = \begin{pmatrix} \hat{A} & \mathbf{a} \\ \mathbf{a}^T & a_{nn} \end{pmatrix}$$

Then as above, $\hat{A}$ is positive definite. Thus it has an LU factorization with all positive entries on the diagonal of U. Notice that, using block multiplication,

$$A = \begin{pmatrix} LU & \mathbf{a} \\ \mathbf{a}^T & a_{nn} \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} U & L^{-1}\mathbf{a} \\ \mathbf{a}^T & a_{nn} \end{pmatrix}$$

Now consider that matrix on the right. Argue that it is of the form $\tilde{L}\tilde{U}$ where $\tilde{U}$ has all positive diagonal entries except possibly for the one in the n^{th} row and n^{th} column. Now explain why $\det(A) > 0$ and argue that in fact all diagonal entries of $\tilde{U}$ are positive.

26. ↑Let A be a real symmetric $n \times n$ matrix and $A = LU$ where L has all ones down the diagonal and U has all positive entries down the main diagonal. Show that $A = LDH$ where L is lower triangular and H is upper triangular, each having all ones down the diagonal and D a diagonal matrix having all positive entries down the main diagonal. In fact, these are the diagonal entries of U.

27. ↑Show that if L, L_1 are lower triangular with ones down the main diagonal and H, H_1 are upper triangular with all ones down the main diagonal and D, D_1 are diagonal matrices having all positive diagonal entries, and if $LDH = L_1 D_1 H_1$, then $L = L_1, H = H_1, D = D_1$. **Hint:** Explain why $D_1^{-1} L_1^{-1} LD = H_1 H^{-1}$. Then explain why the right side is upper triangular and the left side is lower triangular. Conclude these are both diagonal matrices. However, there are all ones down the diagonal in the expression on the right. Hence $H = H_1$. Do something similar to conclude that $L = L_1$ and then that $D = D_1$.

28. ↑Show that if A is a symmetric real matrix such that $\mathbf{x} \to \langle A\mathbf{x}, \mathbf{x}\rangle$ is positive definite, then there exists a lower triangular matrix L having all positive entries down the diagonal such that $A = LL^T$. **Hint:** From the above, $A = LDH$ where L, H are respectively lower and upper triangular having all ones down the diagonal and D is a diagonal matrix having all positive entries. Then argue from the above problem and symmetry of A that $H = L^T$. Now modify L by making it equal to $LD^{1/2}$. This is called the Cholesky factorization.

29. Given $F \in \mathcal{L}(X, Y)$ where X, Y are inner product spaces and $\dim(X) = n \leq m = \dim(Y)$, there exists R, U such that U is nonnegative and Hermitian and $R^* R = I$ such that $F = RU$. Show that U is actually unique and that R is determined on $U(X)$.

INDEX