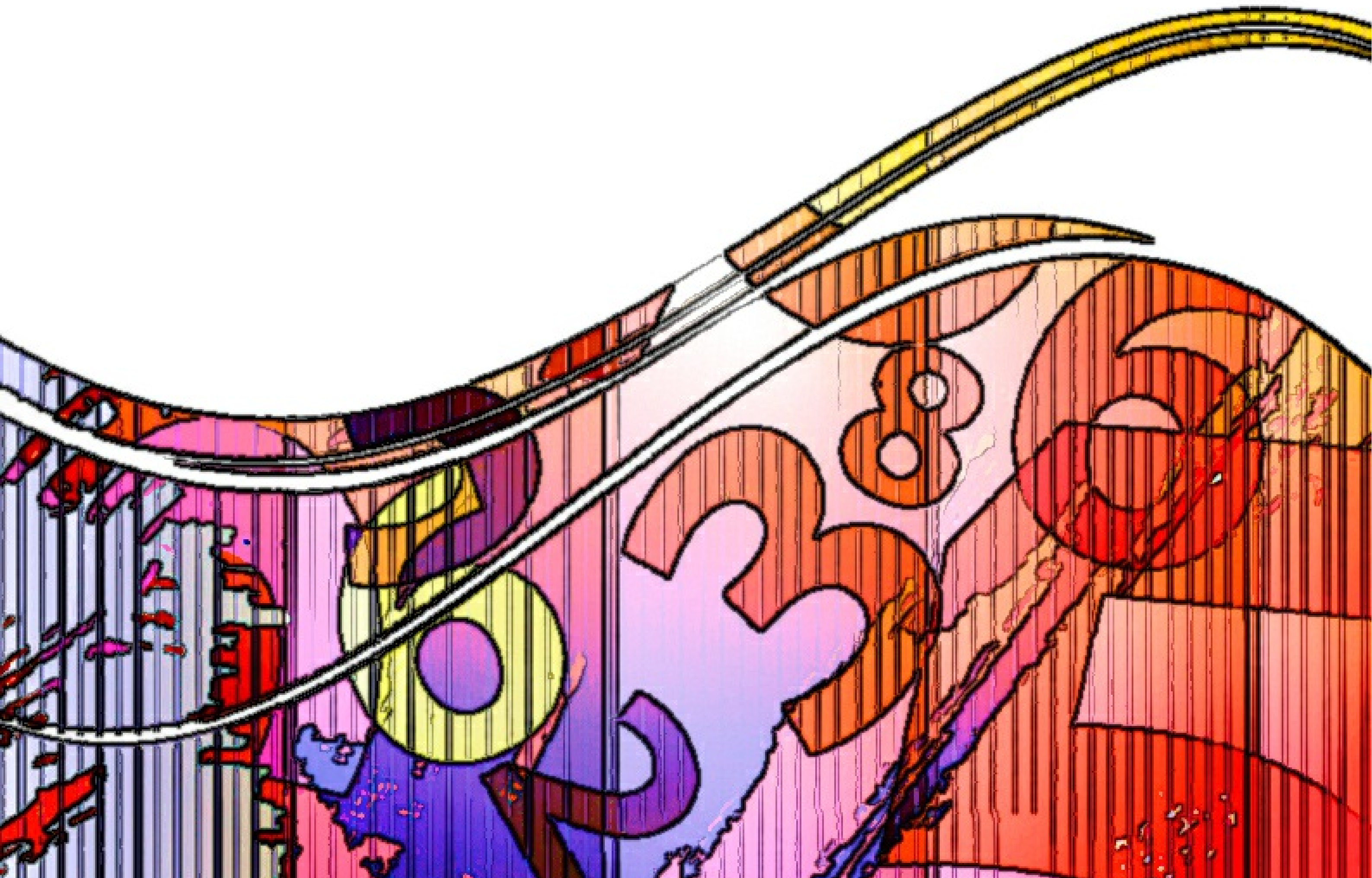
Kenneth Kuttler

Linear Algebra II

Spectral Theory and Abstract Vector Spaces



bookboon.com
The eBook company

Kenneth Kuttler

Linear Algebra II Spectral Theory and Abstract Vector Spaces

Linear Algebra II Spectral Theory and Abstract Vector Spaces

ISBN 978-87-403-0241-7

Contents

To see Chapter 1-6 Download
Linear Algebra I Matrices and Row operations

Spectral Theory

Spectral Theory refers to the study of eigenvalues and eigenvectors of a matrix. It is of fundamental importance in many areas. Row operations will no longer be such a useful tool in this subject.

7.1 Eigenvalues And Eigenvectors Of A Matrix

The field of scalars in spectral theory is best taken to equal $\mathbb{C}$ although I will sometimes refer to it as $\mathbb{F}$ when it could be either $\mathbb{C}$ or $\mathbb{R}$.

Definition 7.1.1 *Let M be an $n \times n$ matrix and let $\mathbf{x} \in \mathbb{C}^n$ be a nonzero vector for which*

$$M\mathbf{x} = \lambda\mathbf{x} \tag{7.1}$$

for some scalar, λ. Then $\mathbf{x}$ is called an eigenvector and λ is called an eigenvalue (characteristic value) of the matrix M.

$\boxed{\textbf{\textit{Eigenvectors are never equal to zero!}}}$

The set of all eigenvalues of an $n \times n$ matrix M, is denoted by $\sigma(M)$ and is referred to as the spectrum of M.

Eigenvectors are vectors which are shrunk, stretched or reflected upon multiplication by a matrix. How can they be identified? Suppose $\mathbf{x}$ satisfies 7.1. Then

$$(\lambda I - M)\mathbf{x} = \mathbf{0}$$

for some $\mathbf{x} \neq \mathbf{0}$. Therefore, the matrix $M - \lambda I$ cannot have an inverse and so by Theorem 3.3.18

$$\det(\lambda I - M) = 0. \tag{7.2}$$

In other words, λ must be a zero of the characteristic polynomial. Since M is an $n \times n$ matrix, it follows from the theorem on expanding a matrix by its cofactor that this is a polynomial equation of degree n. As such, it has a solution, $\lambda \in \mathbb{C}$. Is it actually an eigenvalue? The answer is yes and this follows from Theorem 3.3.26 on Page 123. Since $\det(\lambda I - M) = 0$ the matrix $\lambda I - M$ cannot be one to one and so there exists a nonzero vector, $\mathbf{x}$ such that $(\lambda I - M)\mathbf{x} = \mathbf{0}$. This proves the following corollary.

Corollary 7.1.2 *Let M be an $n \times n$ matrix and* $\det(M - \lambda I) = 0$. *Then there exists* $\mathbf{x} \in \mathbb{C}^n$ *such that* $(M - \lambda I)\mathbf{x} = \mathbf{0}$.

As an example, consider the following.

Example 7.1.3 *Find the eigenvalues and eigenvectors for the matrix*

$$A = \begin{pmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{pmatrix}.$$

You first need to identify the eigenvalues. Recall this requires the solution of the equation

$$\det \left(\lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{pmatrix} \right) = 0$$

When you expand this determinant, you find the equation is

$$(\lambda - 5) \left(\lambda^2 - 20\lambda + 100 \right) = 0$$

and so the eigenvalues are

$$5, 10, 10.$$

I have listed 10 twice because it is a zero of multiplicity two due to

$$\lambda^2 - 20\lambda + 100 = (\lambda - 10)^2 .$$

Having found the eigenvalues, it only remains to find the eigenvectors. First find the eigenvectors for $\lambda = 5$. As explained above, this requires you to solve the equation,

$$\left(5 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

That is you need to find the solution to

$$\begin{pmatrix} 0 & 10 & 5 \\ -2 & -9 & -2 \\ 4 & 8 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

By now this is an old problem. You set up the augmented matrix and row reduce to get the solution. Thus the matrix you must row reduce is

$$\begin{pmatrix} 0 & 10 & 5 & 0 \\ -2 & -9 & -2 & 0 \\ 4 & 8 & -1 & 0 \end{pmatrix}. \tag{7.3}$$

The reduced row echelon form is

$$\begin{pmatrix} 1 & 0 & -\frac{5}{4} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and so the solution is any vector of the form

$$\begin{pmatrix} \frac{5}{4} z \\ \frac{-1}{2} z \\ z \end{pmatrix} = z \begin{pmatrix} \frac{5}{4} \\ \frac{-1}{2} \\ 1 \end{pmatrix}$$

where $z \in \mathbb{F}$. You would obtain the same collection of vectors if you replaced z with $4z$. Thus a simpler description for the solutions to this system of equations whose augmented matrix is in 7.3 is

$$z \begin{pmatrix} 5 \\ -2 \\ 4 \end{pmatrix} \tag{7.4}$$

where $z \in \mathbb{F}$. Now you need to remember that you can't take $z = 0$ because this would result in the zero vector and

$$\boxed{\textbf{\underline{Eigenvectors} \underline{are} \underline{never} \underline{equal} \underline{to} \underline{zero}!}}$$

Other than this value, every other choice of z in 7.4 results in an eigenvector. It is a good idea to check your work! To do so, I will take the original matrix and multiply by this vector and see if I get 5 times this vector.

$$\begin{pmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 25 \\ -10 \\ 20 \end{pmatrix} = 5 \begin{pmatrix} 5 \\ -2 \\ 4 \end{pmatrix}$$

so it appears this is correct. Always check your work on these problems if you care about getting the answer right.

The variable, z is called a free variable or sometimes a parameter. The set of vectors in 7.4 is called the eigenspace and it equals $\ker(\lambda I - A)$. You should observe that in this case the eigenspace has dimension 1 because there is one vector which spans the eigenspace. In general, you obtain the solution from the row echelon form and the number of different free variables gives you the dimension of the eigenspace. Just remember that not every vector in the eigenspace is an eigenvector. The vector, $\mathbf{0}$ is not an eigenvector although it is in the eigenspace because

$$\boxed{\textbf{\underline{Eigenvectors} \underline{are} \underline{never} \underline{equal} \underline{to} \underline{zero}!}}$$

Next consider the eigenvectors for $\lambda = 10$. These vectors are solutions to the equation,

$$\left(10 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

That is you must find the solutions to

$$\begin{pmatrix} 5 & 10 & 5 \\ -2 & -4 & -2 \\ 4 & 8 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which reduces to consideration of the augmented matrix

$$\begin{pmatrix} 5 & 10 & 5 & 0 \\ -2 & -4 & -2 & 0 \\ 4 & 8 & 4 & 0 \end{pmatrix}$$

The row reduced echelon form for this matrix is

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and so the eigenvectors are of the form

$$\begin{pmatrix} -2y-z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

You can't pick z and y both equal to zero because this would result in the zero vector and

Eigenvectors are never equal to zero!

However, every other choice of z and y does result in an eigenvector for the eigenvalue $\lambda = 10$. As in the case for $\lambda = 5$ you should check your work if you care about getting it right.

$$\begin{pmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -10 \\ 0 \\ 10 \end{pmatrix} = 10 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

so it worked. The other vector will also work. Check it.

The above example shows how to find eigenvectors and eigenvalues algebraically. You may have noticed it is a bit long. Sometimes students try to first row reduce the matrix before looking for eigenvalues. This is a terrible idea because row operations destroy the value of the eigenvalues. The eigenvalue problem is really not about row operations. A general rule to remember about the eigenvalue problem is this.

If it is not long and hard it is usually wrong!

The eigenvalue problem is the hardest problem in algebra and people still do research on ways to find eigenvalues. Now if you are so fortunate as to find the eigenvalues as in the above example, then finding the eigenvectors does reduce to row operations and this part of the problem is easy. However, finding the eigenvalues is anything but easy because for an $n \times n$ matrix, it involves solving a polynomial equation of degree n and none of us are very good at doing this. If you only find a good approximation to the eigenvalue, it won't work. It either is or is not an eigenvalue and if it is not, the only solution to the equation, $(\lambda I - M)\,\mathbf{x} = \mathbf{0}$ will be the zero solution as explained above and

$$\boxed{\textbf{Eigenvectors are never equal to zero!}}$$

Here is another example.

Example 7.1.4 *Let*

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

First find the eigenvalues.

$$\det \left(\lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 2 & -2 \\ 1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix} \right) = 0$$

This is $\lambda^3 - 6\lambda^2 + 8\lambda = 0$ and the solutions are 0, 2, and 4.

$$\boxed{0 \textbf{ Can be an Eigenvalue!}}$$

Now find the eigenvectors. For $\lambda = 0$ the augmented matrix for finding the solutions is

$$\begin{pmatrix} 2 & 2 & -2 & 0 \\ 1 & 3 & -1 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix}$$

and the row reduced echelon form is

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

Therefore, the eigenvectors are of the form

$$z \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

where $z \neq 0$.

Next find the eigenvectors for $\lambda = 2$. The augmented matrix for the system of equations needed to find these eigenvectors is

$$\left(\begin{array}{cccc} 0 & -2 & 2 & 0 \\ -1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right)$$

and the row reduced echelon form is

$$\left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

and so the eigenvectors are of the form

$$z\left(\begin{array}{c} 0 \\ 1 \\ 1 \end{array}\right)$$

where $z \neq 0$.

Finally find the eigenvectors for $\lambda = 4$. The augmented matrix for the system of equations needed to find these eigenvectors is

$$\left(\begin{array}{cccc} 2 & -2 & 2 & 0 \\ -1 & 1 & 1 & 0 \\ 1 & -1 & 3 & 0 \end{array}\right)$$

and the row reduced echelon form is

$$\left(\begin{array}{cccc} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Therefore, the eigenvectors are of the form

$$y\left(\begin{array}{c} 1 \\ 1 \\ 0 \end{array}\right)$$

where $y \neq 0$.

Example 7.1.5 *Let*

$$A = \left(\begin{array}{ccc} 2 & -2 & -1 \\ -2 & -1 & -2 \\ 14 & 25 & 14 \end{array}\right).$$

Find the eigenvectors and eigenvalues.

In this case the eigenvalues are $3, 6, 6$ where I have listed 6 twice because it is a zero of algebraic multiplicity two, the characteristic equation being

$$(\lambda - 3)(\lambda - 6)^2 = 0.$$

It remains to find the eigenvectors for these eigenvalues. First consider the eigenvectors for $\lambda = 3$. You must solve

$$\left(3\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) - \left(\begin{array}{ccc} 2 & -2 & -1 \\ -2 & -1 & -2 \\ 14 & 25 & 14 \end{array}\right)\right)\left(\begin{array}{c} x \\ y \\ z \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right).$$

Using routine row operations, the eigenvectors are nonzero vectors of the form

$$\begin{pmatrix} z \\ -z \\ z \end{pmatrix} = z \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Next consider the eigenvectors for $\lambda = 6$. This requires you to solve

$$\left(6\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & -2 & -1 \\ -2 & -1 & -2 \\ 14 & 25 & 14 \end{pmatrix}\right)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and using the usual procedures yields the eigenvectors for $\lambda = 6$ are of the form

$$z\begin{pmatrix} -\frac{1}{8} \\ -\frac{1}{4} \\ 1 \end{pmatrix}$$

or written more simply,

$$z\begin{pmatrix} -1 \\ -2 \\ 8 \end{pmatrix}$$

where $z \in \mathbb{F}$.

Note that in this example the eigenspace for the eigenvalue $\lambda = 6$ is of dimension 1 because there is only one parameter which can be chosen. However, this eigenvalue is of multiplicity two as a root to the characteristic equation.

Definition 7.1.6 *If A is an $n \times n$ matrix with the property that some eigenvalue has algebraic multiplicity as a root of the characteristic equation which is greater than the dimension of the eigenspace associated with this eigenvalue, then the matrix is called defective.*

There may be repeated roots to the characteristic equation, 7.2 and it is not known whether the dimension of the eigenspace equals the multiplicity of the eigenvalue. However, the following theorem is available.

Theorem 7.1.7 *Suppose* $M\mathbf{v}_i = \lambda_i \mathbf{v}_i, i = 1, \cdots, r$, $\mathbf{v}_i \neq 0$, *and that if* $i \neq j$, *then* $\lambda_i \neq \lambda_j$. *Then the set of eigenvectors,* $\{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ *is linearly independent.*

Proof. Suppose the claim of the lemma is not true. Then there exists a subset of this set of vectors

$$\{\mathbf{w}_1, \cdots, \mathbf{w}_r\} \subseteq \{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$$

such that

$$\sum_{j=1}^{r} c_j \mathbf{w}_j = \mathbf{0} \tag{7.5}$$

where each $c_j \neq 0$. Say $M\mathbf{w}_j = \mu_j \mathbf{w}_j$ where

$$\{\mu_1, \cdots, \mu_r\} \subseteq \{\lambda_1, \cdots, \lambda_k\},$$

the μ_j being distinct eigenvalues of M. Out of all such subsets, let this one be such that r is as small as possible. Then necessarily, $r > 1$ because otherwise, $c_1\mathbf{w}_1 = \mathbf{0}$ which would imply $\mathbf{w}_1 = \mathbf{0}$, which is not allowed for eigenvectors.

Now apply M to both sides of 7.5.

$$\sum_{j=1}^{r} c_j \mu_j \mathbf{w}_j = \mathbf{0}. \tag{7.6}$$

Next pick $\mu_k \neq 0$ and multiply both sides of 7.5 by μ_k. Such a μ_k exists because $r > 1$. Thus

$$\sum_{j=1}^{r} c_j \mu_k \mathbf{w}_j = \mathbf{0} \tag{7.7}$$

Subtract the sum in 7.7 from the sum in 7.6 to obtain

$$\sum_{j=1}^{r} c_j \left(\mu_k - \mu_j\right) \mathbf{w}_j = \mathbf{0}$$

Now one of the constants $c_j \left(\mu_k - \mu_j\right)$ equals 0, when $j = k$. Therefore, r was not as small as possible after all. ■

In words, this theorem says that eigenvectors associated with distinct eigenvalues are linearly independent.

Sometimes you have to consider eigenvalues which are complex numbers. This occurs in differential equations for example. You do these problems exactly the same way as you do the ones in which the eigenvalues are real. Here is an example.

Example 7.1.8 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}.$$

You need to find the eigenvalues. Solve

$$\det \left(\lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix} \right) = 0.$$

This reduces to $(\lambda - 1)\left(\lambda^2 - 4\lambda + 5\right) = 0$. The solutions are $\lambda = 1, \lambda = 2 + i, \lambda = 2 - i$.

There is nothing new about finding the eigenvectors for $\lambda = 1$ so consider the eigenvalue $\lambda = 2 + i$. You need to solve

$$\left((2+i) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

In other words, you must consider the augmented matrix

$$\begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & i & 1 & 0 \\ 0 & -1 & i & 0 \end{pmatrix}$$

for the solution. Divide the top row by $(1+i)$ and then take $-i$ times the second row and add to the bottom. This yields

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Now multiply the second row by $-i$ to obtain

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

Therefore, the eigenvectors are of the form

$$z \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}.$$

You should find the eigenvectors for $\lambda = 2 - i$. These are

$$z \begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}.$$

As usual, if you want to get it right you had better check it.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1-2i \\ 2-i \end{pmatrix} = (2-i) \begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}$$

so it worked.

7.2 Some Applications Of Eigenvalues And Eigenvectors

Recall that $n \times n$ matrices can be considered as linear transformations. If F is a 3×3 real matrix having positive determinant, it can be shown that $F = RU$ where R is a rotation matrix and U is a symmetric real matrix having positive eigenvalues. An application of this wonderful result, known to mathematicians as the right polar decomposition, is to continuum mechanics where a chunk of material is identified with a set of points in three dimensional space.

The linear transformation, F in this context is called the deformation gradient and it describes the local deformation of the material. Thus it is possible to consider this deformation in terms of two processes, one which distorts the material and the other which just rotates it. It is the matrix U which is responsible for stretching and compressing. This is why in continuum mechanics, the stress is often taken to depend on U which is known in this context as the right Cauchy Green strain tensor. This process of writing a matrix as a product of two such matrices, one of which preserves distance and the other which distorts is also important in applications to geometric measure theory an interesting field of study in mathematics and to the study of quadratic forms which occur in many applications such as statistics. Here I am emphasizing the application to mechanics in which the eigenvectors of U determine the principle directions, those directions in which the material is stretched or compressed to the maximum extent.

Example 7.2.1 *Find the principle directions determined by the matrix*

$$\begin{pmatrix} \frac{29}{11} & \frac{6}{11} & \frac{6}{11} \\ \frac{6}{11} & \frac{41}{44} & \frac{19}{44} \\ \frac{6}{11} & \frac{19}{44} & \frac{41}{44} \end{pmatrix}$$

The eigenvalues are $3, 1,$ *and* $\frac{1}{2}$.

It is nice to be given the eigenvalues. The largest eigenvalue is 3 which means that in the direction determined by the eigenvector associated with 3 the stretch is three times as large. The smallest eigenvalue is $1/2$ and so in the direction determined by the eigenvector for $1/2$ the material is compressed, becoming locally half as long. It remains to find these directions. First consider the eigenvector for 3. It is necessary to solve

$$\left(3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{29}{11} & \frac{6}{11} & \frac{6}{11} \\ \frac{6}{11} & \frac{41}{44} & \frac{19}{44} \\ \frac{6}{11} & \frac{19}{44} & \frac{41}{44} \end{pmatrix}\right)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Thus the augmented matrix for this system of equations is

$$\begin{pmatrix} \frac{4}{11} & -\frac{6}{11} & -\frac{6}{11} & 0 \\ -\frac{6}{11} & \frac{91}{44} & -\frac{19}{44} & 0 \\ -\frac{6}{11} & -\frac{19}{44} & \frac{91}{44} & 0 \end{pmatrix}$$

The row reduced echelon form is

$$\begin{pmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and so the principle direction for the eigenvalue 3 in which the material is stretched to the maximum extent is

$$\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

A direction vector in this direction is

$$\begin{pmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{pmatrix}.$$

You should show that the direction in which the material is compressed the most is in the direction

$$\begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Note this is meaningful information which you would have a hard time finding without the theory of eigenvectors and eigenvalues.

Another application is to the problem of finding solutions to systems of differential equations. It turns out that vibrating systems involving masses and springs can be studied in the form

$$\mathbf{x}'' = A\mathbf{x} \tag{7.8}$$

where A is a real symmetric $n \times n$ matrix which has nonpositive eigenvalues. This is analogous to the case of the scalar equation for undamped oscillation, $x'' + \omega^2 x = 0$. The main difference is that here the scalar ω^2 is replaced with the matrix $-A$. Consider the problem of finding solutions to 7.8. You look for a solution which is in the form

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t} \tag{7.9}$$

and substitute this into 7.8. Thus

$$\mathbf{x}'' = \mathbf{v}\lambda^2 e^{\lambda t} = e^{\lambda t} A\mathbf{v}$$

and so

$$\lambda^2 \mathbf{v} = A\mathbf{v}.$$

Therefore, λ^2 needs to be an eigenvalue of A and $\mathbf{v}$ needs to be an eigenvector. Since A has nonpositive eigenvalues, $\lambda^2 = -a^2$ and so $\lambda = \pm ia$ where $-a^2$ is an eigenvalue of A. Corresponding to this you obtain solutions of the form

$$\mathbf{x}(t) = \mathbf{v}\cos(at), \mathbf{v}\sin(at).$$

Note these solutions oscillate because of the $\cos(at)$ and $\sin(at)$ in the solutions. Here is an example.

Example 7.2.2 *Find oscillatory solutions to the system of differential equations,* $\mathbf{x}'' = A\mathbf{x}$ *where*

$$A = \begin{pmatrix} -\frac{5}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{13}{6} & \frac{5}{6} \\ -\frac{1}{3} & \frac{5}{6} & -\frac{13}{6} \end{pmatrix}.$$

The eigenvalues are $-1, -2,$ *and* -3.

According to the above, you can find solutions by looking for the eigenvectors. Consider the eigenvectors for -3. The augmented matrix for finding the eigenvectors is

$$\begin{pmatrix} -\frac{4}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & -\frac{5}{6} & -\frac{5}{6} & 0 \\ \frac{1}{3} & -\frac{5}{6} & -\frac{5}{6} & 0 \end{pmatrix}$$

and its row echelon form is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore, the eigenvectors are of the form

$$\mathbf{v} = z \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

It follows

$$\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cos\left(\sqrt{3}t\right), \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \sin\left(\sqrt{3}t\right)$$

are both solutions to the system of differential equations. You can find other oscillatory solutions in the same way by considering the other eigenvalues. You might try checking these answers to verify they work.

This is just a special case of a procedure used in differential equations to obtain closed form solutions to systems of differential equations using linear algebra. The overall philosophy is to take one of the easiest problems in analysis and change it into the eigenvalue problem which is the most difficult problem in algebra. However, when it works, it gives precise solutions in terms of known functions.

7.3 Exercises

1. If A is the matrix of a linear transformation which rotates all vectors in $\mathbb{R}^2$ through 30°, explain why A cannot have any real eigenvalues.

2. If A is an $n \times n$ matrix and c is a nonzero constant, compare the eigenvalues of A and cA.

3. If A is an invertible $n \times n$ matrix, compare the eigenvalues of A and A^{-1}. More generally, for m an arbitrary integer, compare the eigenvalues of A and A^m.

4. Let A, B be invertible $n \times n$ matrices which commute. That is, $AB = BA$. Suppose $\mathbf{x}$ is an eigenvector of B. Show that then $A\mathbf{x}$ must also be an eigenvector for B.

5. Suppose A is an $n \times n$ matrix and it satisfies $A^m = A$ for some m a positive integer larger than 1. Show that if λ is an eigenvalue of A then $|\lambda|$ equals either 0 or 1.

6. Show that if $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \lambda\mathbf{y}$, then whenever a, b are scalars,
$$A(a\mathbf{x} + b\mathbf{y}) = \lambda(a\mathbf{x} + b\mathbf{y}).$$
Does this imply that $a\mathbf{x} + b\mathbf{y}$ is an eigenvector? Explain.

7. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} -1 & -1 & 7 \\ -1 & 0 & 4 \\ -1 & -1 & 5 \end{pmatrix}$. Determine whether the matrix is defective.

8. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} -3 & -7 & 19 \\ -2 & -1 & 8 \\ -2 & -3 & 10 \end{pmatrix}$.Determine whether the matrix is defective.

9. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} -7 & -12 & 30 \\ -3 & -7 & 15 \\ -3 & -6 & 14 \end{pmatrix}$.

10. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 7 & -2 & 0 \\ 8 & -1 & 0 \\ -2 & 4 & 6 \end{pmatrix}$. Determine whether the matrix is defective.
11. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 3 & -2 & -1 \\ 0 & 5 & 1 \\ 0 & 2 & 4 \end{pmatrix}$.
12. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 6 & 8 & -23 \\ 4 & 5 & -16 \\ 3 & 4 & -12 \end{pmatrix}$. Determine whether the matrix is defective.
13. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 5 & 2 & -5 \\ 12 & 3 & -10 \\ 12 & 4 & -11 \end{pmatrix}$. Determine whether the matrix is defective.

14. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 20 & 9 & -18 \\ 6 & 5 & -6 \\ 30 & 14 & -27 \end{pmatrix}$. Determine whether the matrix is defective.

15. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 1 & 26 & -17 \\ 4 & -4 & 4 \\ -9 & -18 & 9 \end{pmatrix}$. Determine whether the matrix is defective.

16. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 3 & -1 & -2 \\ 11 & 3 & -9 \\ 8 & 0 & -6 \end{pmatrix}$. Determine whether the matrix is defective.

17. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} -2 & 1 & 2 \\ -11 & -2 & 9 \\ -8 & 0 & 7 \end{pmatrix}$. Determine whether the matrix is defective.

18. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 2 & 1 & -1 \\ 2 & 3 & -2 \\ 2 & 2 & -1 \end{pmatrix}$. Determine whether the matrix is defective.

19. Find the complex eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 4 & -2 & -2 \\ 0 & 2 & -2 \\ 2 & 0 & 2 \end{pmatrix}$.

20. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 9 & 6 & -3 \\ 0 & 6 & 0 \\ -3 & -6 & 9 \end{pmatrix}$. Determine whether the matrix is defective.

21. Find the complex eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 4 & -2 & -2 \\ 0 & 2 & -2 \\ 2 & 0 & 2 \end{pmatrix}$. Determine whether the matrix is defective.

22. Find the complex eigenvalues and eigenvectors of the matrix $\begin{pmatrix} -4 & 2 & 0 \\ 2 & -4 & 0 \\ -2 & 2 & -2 \end{pmatrix}$. Determine whether the matrix is defective.

23. Find the complex eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 1 & 1 & -6 \\ 7 & -5 & -6 \\ -1 & 7 & 2 \end{pmatrix}$. Determine whether the matrix is defective.

24. Find the complex eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 4 & 2 & 0 \\ -2 & 4 & 0 \\ -2 & 2 & 6 \end{pmatrix}$. Determine whether the matrix is defective.

25. Here is a matrix.

$$\begin{pmatrix} 1 & a & 0 & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 2 & c \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Find values of a, b, c for which the matrix is defective and values of a, b, c for which it is nondefective.

26. Here is a matrix.
$$\begin{pmatrix} a & 1 & 0 \\ 0 & b & 1 \\ 0 & 0 & c \end{pmatrix}$$
where a, b, c are numbers. Show this is sometimes defective depending on the choice of a, b, c. What is an easy case which will ensure it is not defective?

27. Suppose A is an $n \times n$ matrix consisting entirely of real entries but $a + ib$ is a complex eigenvalue having the eigenvector, $\mathbf{x} + i\mathbf{y}$. Here $\mathbf{x}$ and $\mathbf{y}$ are real vectors. Show that then $a - ib$ is also an eigenvalue with the eigenvector, $\mathbf{x} - i\mathbf{y}$. **Hint:** You should remember that the conjugate of a product of complex numbers equals the product of the conjugates. Here $a + ib$ is a complex number whose conjugate equals $a - ib$.

28. Recall an $n \times n$ matrix is said to be symmetric if it has all real entries and if $A = A^T$. Show the eigenvalues of a real symmetric matrix are real and for each eigenvalue, it has a real eigenvector.

29. Recall an $n \times n$ matrix is said to be skew symmetric if it has all real entries and if $A = -A^T$. Show that any nonzero eigenvalues must be of the form ib where $i^2 = -1$. In words, the eigenvalues are either 0 or pure imaginary.

30. Is it possible for a nonzero matrix to have only 0 as an eigenvalue?

31. Show that the eigenvalues and eigenvectors of a real matrix occur in conjugate pairs.

32. Suppose A is an $n \times n$ matrix having all real eigenvalues which are distinct. Show there exists S such that $S^{-1}AS = D$, a diagonal matrix. If
$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
define e^D by
$$e^D \equiv \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}$$
and define
$$e^A \equiv Se^D S^{-1}.$$
Next show that if A is as just described, so is tA where t is a real number and the eigenvalues of At are $t\lambda_k$. If you differentiate a matrix of functions entry by entry so that for the ij^{th} entry of $A'(t)$ you get $a'_{ij}(t)$ where $a_{ij}(t)$ is the ij^{th} entry of $A(t)$, show
$$\frac{d}{dt}\left(e^{At}\right) = Ae^{At}$$
Next show $\det\left(e^{At}\right) \neq 0$. This is called the matrix exponential. Note I have only defined it for the case where the eigenvalues of A are real, but the same procedure will work even for complex eigenvalues. All you have to do is to define what is meant by e^{a+ib}.

33. Find the principle directions determined by the matrix $\begin{pmatrix} \frac{7}{12} & -\frac{1}{4} & \frac{1}{6} \\ -\frac{1}{4} & \frac{7}{12} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & \frac{2}{3} \end{pmatrix}$. The eigenvalues are $\frac{1}{3}, 1$, and $\frac{1}{2}$ listed according to multiplicity.

34. Find the principle directions determined by the matrix $\begin{pmatrix} \frac{5}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{7}{6} & \frac{1}{6} \\ -\frac{1}{3} & \frac{1}{6} & \frac{7}{6} \end{pmatrix}$ The eigenvalues are $1, 2$, and 1. What is the physical interpretation of the repeated eigenvalue?

35. Find oscillatory solutions to the system of differential equations, $\mathbf{x}'' = A\mathbf{x}$ where $A = \begin{pmatrix} -3 & -1 & -1 \\ -1 & -2 & 0 \\ -1 & 0 & -2 \end{pmatrix}$ The eigenvalues are $-1, -4$, and -2.

36. Let A and B be $n \times n$ matrices and let the columns of B be

$$\mathbf{b}_1, \cdots, \mathbf{b}_n$$

and the rows of A are

$$\mathbf{a}_1^T, \cdots, \mathbf{a}_n^T.$$

Show the columns of AB are

$$A\mathbf{b}_1 \cdots A\mathbf{b}_n$$

and the rows of AB are

$$\mathbf{a}_1^T B \cdots \mathbf{a}_n^T B.$$

37. Let M be an $n \times n$ matrix. Then define the adjoint of M, denoted by M^* to be the transpose of the conjugate of M. For example,

$$\begin{pmatrix} 2 & i \\ 1+i & 3 \end{pmatrix}^* = \begin{pmatrix} 2 & 1-i \\ -i & 3 \end{pmatrix}.$$

A matrix M, is self adjoint if $M^* = M$. Show the eigenvalues of a self adjoint matrix are all real.

38. Let M be an $n \times n$ matrix and suppose $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are n eigenvectors which form a linearly independent set. Form the matrix S by making the columns these vectors. Show that S^{-1} exists and that $S^{-1}MS$ is a diagonal matrix (one having zeros everywhere except on the main diagonal) having the eigenvalues of M on the main diagonal. When this can be done the matrix is said to be diagonalizable.

39. Show that a $n \times n$ matrix M is diagonalizable if and only if $\mathbb{F}^n$ has a basis of eigenvectors. **Hint:** The first part is done in Problem 38. It only remains to show that if the

matrix can be diagonalized by some matrix S giving $D = S^{-1}MS$ for D a diagonal matrix, then it has a basis of eigenvectors. Try using the columns of the matrix S.

40. Let

$$A = \begin{pmatrix} \boxed{\begin{matrix} 1 & 2 \\ 3 & 4 \end{matrix}} & \boxed{\begin{matrix} 2 \\ 0 \end{matrix}} \\ \boxed{\begin{matrix} 0 & 1 \end{matrix}} & \boxed{3} \end{pmatrix}$$

and let

$$B = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ 1 & 1 \end{matrix}} \\ \boxed{\begin{matrix} 2 & 1 \end{matrix}} \end{pmatrix}$$

Multiply AB verifying the block multiplication formula. Here $A_{11} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $A_{12} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $A_{21} = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and $A_{22} = (3)$.

41. Suppose A, B are $n \times n$ matrices and λ is a nonzero eigenvalue of AB. Show that then it is also an eigenvalue of BA. **Hint:** Use the definition of what it means for λ to be an eigenvalue. That is,

$$AB\mathbf{x} = \lambda \mathbf{x}$$

where $\mathbf{x} \neq \mathbf{0}$. Maybe you should multiply both sides by B.

42. Using the above problem show that if A, B are $n \times n$ matrices, it is not possible that $AB - BA = aI$ for any $a \neq 0$. **Hint:** First show that if A is a matrix, then the eigenvalues of $A - aI$ are $\lambda - a$ where λ is an eigenvalue of A.

43. Consider the following matrix.

$$C = \begin{pmatrix} 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & -a_1 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & -a_{n-1} \end{pmatrix}$$

Show $\det(\lambda I - C) = a_0 + \lambda a_1 + \cdots a_{n-1}\lambda^{n-1} + \lambda^n$. This matrix is called a companion matrix for the given polynomial.

44. A discreet dynamical system is of the form

$$\mathbf{x}(k+1) = A\mathbf{x}(k), \ \mathbf{x}(0) = \mathbf{x}_0$$

where A is an $n \times n$ matrix and $\mathbf{x}(k)$ is a vector in $\mathbb{R}^n$. Show first that

$$\mathbf{x}(k) = A^k \mathbf{x}_0$$

for all $k \geq 1$. If A is nondefective so that it has a basis of eigenvectors, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ where

$$A\mathbf{v}_j = \lambda_j \mathbf{v}_j$$

you can write the initial condition $\mathbf{x}_0$ in a unique way as a linear combination of these eigenvectors. Thus

$$\mathbf{x}_0 = \sum_{j=1}^{n} a_j \mathbf{v}_j$$

Now explain why

$$\mathbf{x}(k) = \sum_{j=1}^{n} a_j A^k \mathbf{v}_j = \sum_{j=1}^{n} a_j \lambda_j^k \mathbf{v}_j$$

which gives a formula for $\mathbf{x}(k)$, the solution of the dynamical system.

45. Suppose A is an $n \times n$ matrix and let $\mathbf{v}$ be an eigenvector such that $A\mathbf{v} = \lambda \mathbf{v}$. Also suppose the characteristic polynomial of A is

$$\det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

Explain why

$$\left(A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I\right)\mathbf{v} = \mathbf{0}$$

If A is nondefective, give a very easy proof of the Cayley Hamilton theorem based on this. Recall this theorem says A satisfies its characteristic equation,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0.$$

46. Suppose an $n \times n$ nondefective matrix A has only 1 and -1 as eigenvalues. Find A^{12}.

47. Suppose the characteristic polynomial of an $n \times n$ matrix A is $1 - \lambda^n$. Find A^{mn} where m is an integer. **Hint:** Note first that A is nondefective. Why?

48. Sometimes sequences come in terms of a recursion formula. An example is the Fibonacci sequence.

$$x_0 = 1 = x_1,\ x_{n+1} = x_n + x_{n-1}$$

Show this can be considered as a discreet dynamical system as follows.

$$\begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_n \\ x_{n-1} \end{pmatrix}, \begin{pmatrix} x_1 \\ x_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Now use the technique of Problem 44 to find a formula for x_n.

49. Let A be an $n \times n$ matrix having characteristic polynomial

$$\det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

Show that $a_0 = (-1)^n \det(A)$.

7.4 Schur's Theorem

Every matrix is related to an upper triangular matrix in a particularly significant way. This is Schur's theorem and it is the most important theorem in the spectral theory of matrices.

Lemma 7.4.1 *Let $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ be a basis for $\mathbb{F}^n$. Then there exists an orthonormal basis for $\mathbb{F}^n$, $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ which has the property that for each $k \leq n$, span$(\mathbf{x}_1, \cdots, \mathbf{x}_k) =$* span$(\mathbf{u}_1, \cdots, \mathbf{u}_k)$.

Proof: Let $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ be a basis for $\mathbb{F}^n$. Let $\mathbf{u}_1 \equiv \mathbf{x}_1 / |\mathbf{x}_1|$. Thus for $k = 1$, span$(\mathbf{u}_1) =$ span$(\mathbf{x}_1)$ and $\{\mathbf{u}_1\}$ is an orthonormal set. Now suppose for some $k < n$, $\mathbf{u}_1, \cdots, \mathbf{u}_k$ have been chosen such that $(\mathbf{u}_j \cdot \mathbf{u}_l) = \delta_{jl}$ and span$(\mathbf{x}_1, \cdots, \mathbf{x}_k) =$ span$(\mathbf{u}_1, \cdots, \mathbf{u}_k)$. Then define

$$\mathbf{u}_{k+1} \equiv \frac{\mathbf{x}_{k+1} - \sum_{j=1}^{k} (\mathbf{x}_{k+1} \cdot \mathbf{u}_j)\mathbf{u}_j}{\left|\mathbf{x}_{k+1} - \sum_{j=1}^{k} (\mathbf{x}_{k+1} \cdot \mathbf{u}_j)\mathbf{u}_j\right|}, \tag{7.10}$$

where the denominator is not equal to zero because the $\mathbf{x}_j$ form a basis and so

$$\mathbf{x}_{k+1} \notin \text{span}\left(\mathbf{x}_1, \cdots, \mathbf{x}_k\right) = \text{span}\left(\mathbf{u}_1, \cdots, \mathbf{u}_k\right)$$

Thus by induction,

$$\mathbf{u}_{k+1} \in \text{span}\left(\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{x}_{k+1}\right) = \text{span}\left(\mathbf{x}_1, \cdots, \mathbf{x}_k, \mathbf{x}_{k+1}\right).$$

Also, $\mathbf{x}_{k+1} \in \text{span}\left(\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}\right)$ which is seen easily by solving 7.10 for $\mathbf{x}_{k+1}$ and it follows

$$\text{span}\left(\mathbf{x}_1, \cdots, \mathbf{x}_k, \mathbf{x}_{k+1}\right) = \text{span}\left(\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}\right).$$

If $l \leq k$,

$$\left(\mathbf{u}_{k+1} \cdot \mathbf{u}_l\right) = C\left(\left(\mathbf{x}_{k+1} \cdot \mathbf{u}_l\right) - \sum_{j=1}^{k} \left(\mathbf{x}_{k+1} \cdot \mathbf{u}_j\right)\left(\mathbf{u}_j \cdot \mathbf{u}_l\right)\right) =$$

$$C\left(\left(\mathbf{x}_{k+1} \cdot \mathbf{u}_l\right) - \sum_{j=1}^{k} \left(\mathbf{x}_{k+1} \cdot \mathbf{u}_j\right)\delta_{lj}\right) = C\left(\left(\mathbf{x}_{k+1} \cdot \mathbf{u}_l\right) - \left(\mathbf{x}_{k+1} \cdot \mathbf{u}_l\right)\right) = 0.$$

The vectors, $\{\mathbf{u}_j\}_{j=1}^n$, 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. Here is a fundamental definition.

Definition 7.4.2 *An $n \times n$ matrix U, is unitary if $UU^* = I = U^*U$ where U^* is defined to be the transpose of the conjugate of U.*

Proposition 7.4.3 *An $n \times n$ matrix is unitary if and only if the columns are an orthonormal set.*

Proof: This follows right away from the way we multiply matrices. If U is an $n \times n$ complex matrix, then

$$\left(U^*U\right)_{ij} = \mathbf{u}_i^*\mathbf{u}_j = \overline{\left(\mathbf{u}_i, \mathbf{u}_j\right)}$$

and the matrix is unitary if and only if this equals δ_{ij} if and only if the columns are orthonormal. ■

Theorem 7.4.4 *Let A be an $n \times n$ matrix. Then there exists a unitary matrix U such that*

$$U^*AU = T, \tag{7.11}$$

where T is an upper triangular matrix having the eigenvalues of A on the main diagonal *listed according to multiplicity as roots of the characteristic equation.*

Proof: The theorem is clearly true if A is a 1×1 matrix. Just let $U = 1$ the 1×1 matrix which has 1 down the main diagonal and zeros elsewhere. Suppose it is true for $(n-1) \times (n-1)$ matrices and let A be an $n \times n$ matrix. Then let $\mathbf{v}_1$ be a unit eigenvector for A . Then there exists λ_1 such that

$$A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, \ |\mathbf{v}_1| = 1.$$

Extend $\{\mathbf{v}_1\}$ to a basis and then use Lemma 7.4.1 to obtain $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, an orthonormal basis in $\mathbb{F}^n$. Let U_0 be a matrix whose i^{th} column is $\mathbf{v}_i$. Then from the above, it follows U_0 is unitary. Then $U_0^* A U_0$ is of the form

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

where A_1 is an $n-1 \times n-1$ matrix. Now by induction there exists an $(n-1) \times (n-1)$ unitary matrix $\widetilde{U}_1$ such that

$$\widetilde{U}_1^* A_1 \widetilde{U}_1 = T_{n-1},$$

an upper triangular matrix. Consider

$$U_1 \equiv \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widetilde{U}_1 \end{pmatrix}$$

This is a unitary matrix and

$$U_1^* U_0^* A U_0 U_1 = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widetilde{U}_1^* \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & A_1 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widetilde{U}_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & T_{n-1} \end{pmatrix} \equiv T$$

where T is upper triangular. Then let $U = U_0 U_1$. Since $(U_0 U_1)^* = U_1^* U_0^*$, it follows A is similar to T and that $U_0 U_1$ is unitary. Hence A and T have the same characteristic polynomials and since the eigenvalues of T are the diagonal entries listed according to algebraic multiplicity, ■

As a simple consequence of the above theorem, here is an interesting lemma.

Lemma 7.4.5 *Let A be of the form*

$$A = \begin{pmatrix} P_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & P_s \end{pmatrix}$$

where P_k is an $m_k \times m_k$ matrix. Then

$$\det(A) = \prod_k \det(P_k).$$

Also, the eigenvalues of A consist of the union of the eigenvalues of the P_j.

Proof: Let U_k be an $m_k \times m_k$ unitary matrix such that

$$U_k^* P_k U_k = T_k$$

where T_k is upper triangular. Then it follows that for

$$U \equiv \begin{pmatrix} U_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_s \end{pmatrix}, \ U^* = \begin{pmatrix} U_1^* & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_s^* \end{pmatrix}$$

and also

$$\begin{pmatrix} U_1^* & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_s^* \end{pmatrix} \begin{pmatrix} P_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & P_s \end{pmatrix} \begin{pmatrix} U_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_s \end{pmatrix} = \begin{pmatrix} T_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & T_s \end{pmatrix}.$$

Therefore, since the determinant of an upper triangular matrix is the product of the diagonal entries,

$$\det(A) = \prod_k \det(T_k) = \prod_k \det(P_k).$$

From the above formula, the eigenvalues of A consist of the eigenvalues of the upper triangular matrices T_k, and each T_k has the same eigenvalues as P_k. ■

What if A is a real matrix and you only want to consider real unitary matrices?

Theorem 7.4.6 *Let A be a real $n \times n$ matrix. Then there exists a real unitary matrix Q and a matrix T of the form*

$$T = \begin{pmatrix} P_1 & \cdots & * \\ & \ddots & \vdots \\ 0 & & P_r \end{pmatrix} \tag{7.12}$$

where P_i equals either a real 1×1 matrix or P_i equals a real 2×2 matrix having as its eigenvalues a conjugate pair of eigenvalues of A such that $Q^T AQ = T$. The matrix T is called the real Schur form of the matrix A. Recall that a real unitary matrix is also called an orthogonal matrix.

Proof: Suppose

$$A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, \ |\mathbf{v}_1| = 1$$

where λ_1 is real. Then let $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ be an orthonormal basis of vectors in $\mathbb{R}^n$. Let Q_0 be a matrix whose i^{th} column is $\mathbf{v}_i$. Then $Q_0^* A Q_0$ is of the form

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

where A_1 is a real $n-1 \times n-1$ matrix. This is just like the proof of Theorem 7.4.4 up to this point.

Now consider the case where $\lambda_1 = \alpha + i\beta$ where $\beta \neq 0$. It follows since A is real that $\mathbf{v}_1 = \mathbf{z}_1 + i\mathbf{w}_1$ and that $\overline{\mathbf{v}}_1 = \mathbf{z}_1 - i\mathbf{w}_1$ is an eigenvector for the eigenvalue $\alpha - i\beta$. Here $\mathbf{z}_1$ and $\mathbf{w}_1$ are real vectors. Since $\overline{\mathbf{v}}_1$ and $\mathbf{v}_1$ are eigenvectors corresponding to distinct eigenvalues, they form a linearly independent set. From this it follows that $\{\mathbf{z}_1, \mathbf{w}_1\}$ is an independent set of vectors in $\mathbb{C}^n$, hence in $\mathbb{R}^n$. Indeed,$\{\mathbf{v}_1, \overline{\mathbf{v}}_1\}$ is an independent set and also $\operatorname{span}(\mathbf{v}_1, \overline{\mathbf{v}}_1) = \operatorname{span}(\mathbf{z}_1, \mathbf{w}_1)$. Now using the Gram Schmidt theorem in $\mathbb{R}^n$, there exists $\{\mathbf{u}_1, \mathbf{u}_2\}$, an orthonormal set of real vectors such that $\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2) = \operatorname{span}(\mathbf{v}_1, \overline{\mathbf{v}}_1)$. For example,

$$\mathbf{u}_1 = \mathbf{z}_1 / |\mathbf{z}_1|, \ \mathbf{u}_2 = \frac{|\mathbf{z}_1|^2 \mathbf{w}_1 - (\mathbf{w}_1 \cdot \mathbf{z}_1) \mathbf{z}_1}{\left| |\mathbf{z}_1|^2 \mathbf{w}_1 - (\mathbf{w}_1 \cdot \mathbf{z}_1) \mathbf{z}_1 \right|}$$

Let $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be an orthonormal basis in $\mathbb{R}^n$ and let Q_0 be a unitary matrix whose i^{th} column is $\mathbf{u}_i$ so Q_0 is a real orthogonal matrix. Then $A\mathbf{u}_j$ are both in $\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$ for

$j = 1, 2$ and so $\mathbf{u}_k^T A \mathbf{u}_j = 0$ whenever $k \geq 3$. It follows that $Q_0^* A Q_0$ is of the form

$$Q_0^* A Q_0 = \begin{pmatrix} * & * & \cdots & * \\ * & * & & \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} P_1 & * \\ 0 & A_1 \end{pmatrix}$$

where A_1 is now an $n-2 \times n-2$ matrix and P_1 is a 2×2 matrix. Now this is similar to A and so two of its eigenvalues are $\alpha + i\beta$ and $\alpha - i\beta$.

Now find $\widetilde{Q}_1$ an $n-2 \times n-2$ matrix to put A_1 in an appropriate form as above and come up with A_2 either an $n-4 \times n-4$ matrix or an $n-3 \times n-3$ matrix. Then the only other difference is to let

$$Q_1 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & & & \\ \vdots & \vdots & & \widetilde{Q}_1 & \\ 0 & 0 & & & \end{pmatrix}$$

thus putting a 2×2 identity matrix in the upper left corner rather than a one. Repeating this process with the above modification for the case of a complex eigenvalue leads eventually to 7.12 where Q is the product of real unitary matrices Q_i above. When the block P_i is 2×2, its eigenvalues are a conjugate pair of eigenvalues of A and if it is 1×1 it is a real eigenvalue of A.

Here is why this last claim is true

$$\lambda I - T = \begin{pmatrix} \lambda I_1 - P_1 & \cdots & * \\ & \ddots & \vdots \\ 0 & & \lambda I_r - P_r \end{pmatrix}$$

where I_k is the 2×2 identity matrix in the case that P_k is 2×2 and is the number 1 in the case where P_k is a 1×1 matrix. Now by Lemma 7.4.5,

$$\det(\lambda I - T) = \prod_{k=1}^{r} \det(\lambda I_k - P_k).$$

Therefore, λ is an eigenvalue of T if and only if it is an eigenvalue of some P_k. This proves the theorem since the eigenvalues of T are the same as those of A including multiplicity because they have the same characteristic polynomial due to the similarity of A and T. ■

Corollary 7.4.7 *Let A be a real $n \times n$ matrix having only real eigenvalues. Then there exists a real orthogonal matrix Q and an upper triangular matrix T such that*

$$Q^T A Q = T$$

and furthermore, if the eigenvalues of A are listed in decreasing order,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

Q can be chosen such that T is of the form

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Proof: Most of this follows right away from Theorem 7.4.6. It remains to verify the claim that the diagonal entries can be arranged in the desired order. However, this follows from a simple modification of the above argument. When you find $\mathbf{v}_1$ the eigenvalue of λ_1, just be sure λ_1 is chosen to be the largest eigenvalue. Then in the rest of the argument, always choose the largest eigenvalue at each step of the construction. ■

Of course there is a similar conclusion which can be proved exactly the same way in the case where A has complex eigenvalues.

Corollary 7.4.8 *Let A be a real $n \times n$ matrix. Then there exists a real orthogonal matrix Q and an upper triangular matrix T such that*

$$Q^T A Q = T = \begin{pmatrix} P_1 & \cdots & * \\ & \ddots & \vdots \\ 0 & & P_r \end{pmatrix}$$

where P_i equals either a real 1×1 matrix or P_i equals a real 2×2 matrix having as its eigenvalues a conjugate pair of eigenvalues of A. If P_k corresponds to the two eigenvalues $\alpha_k \pm i\beta_k \equiv \sigma(P_k)$, Q can be chosen such that

$$|\sigma(P_1)| \geq |\sigma(P_2)| \geq \cdots$$

where

$$|\sigma(P_k)| \equiv \sqrt{\alpha_k^2 + \beta_k^2}$$

The blocks, P_k can be arranged in any other order also.

Definition 7.4.9 *When a linear transformation A, mapping a linear space V to V has a basis of eigenvectors, the linear transformation is called non defective. Otherwise it is called defective. An $n \times n$ matrix A, is called normal if $AA^* = A^*A$. An important class of normal matrices is that of the Hermitian or self adjoint matrices. An $n \times n$ matrix A is self adjoint or Hermitian if $A = A^*$.*

You can check that an example of a normal matrix which is neither symmetric nor Hermitian is $\begin{pmatrix} 6i & -(1+i)\sqrt{2} \\ (1-i)\sqrt{2} & 6i \end{pmatrix}$.

The next lemma is the basis for concluding that every normal matrix is unitarily similar to a diagonal matrix.

Lemma 7.4.10 *If T is upper triangular and normal, then T is a diagonal matrix.*

Proof: This is obviously true if T is 1×1. In fact, it can't help being diagonal in this case. Suppose then that the lemma is true for $(n-1) \times (n-1)$ matrices and let T be an upper triangular normal $n \times n$ matrix. Thus T is of the form

$$T = \begin{pmatrix} t_{11} & \mathbf{a}^* \\ \mathbf{0} & T_1 \end{pmatrix},\ T^* = \begin{pmatrix} \overline{t_{11}} & \mathbf{0}^T \\ \mathbf{a} & T_1^* \end{pmatrix}$$

Then

$$\begin{aligned} TT^* &= \begin{pmatrix} t_{11} & \mathbf{a}^* \\ \mathbf{0} & T_1 \end{pmatrix} \begin{pmatrix} \overline{t_{11}} & \mathbf{0}^T \\ \mathbf{a} & T_1^* \end{pmatrix} = \begin{pmatrix} |t_{11}|^2 + \mathbf{a}^*\mathbf{a} & \mathbf{a}^* T_1^* \\ T_1 \mathbf{a} & T_1 T_1^* \end{pmatrix} \\ T^*T &= \begin{pmatrix} \overline{t_{11}} & \mathbf{0}^T \\ \mathbf{a} & T_1^* \end{pmatrix} \begin{pmatrix} t_{11} & \mathbf{a}^* \\ \mathbf{0} & T_1 \end{pmatrix} = \begin{pmatrix} |t_{11}|^2 & \overline{t_{11}}\mathbf{a}^* \\ \mathbf{a} t_{11} & \mathbf{a}\mathbf{a}^* + T_1^* T_1 \end{pmatrix} \end{aligned}$$

Since these two matrices are equal, it follows $\mathbf{a} = \mathbf{0}$. But now it follows that $T_1^* T_1 = T_1 T_1^*$ and so by induction T_1 is a diagonal matrix D_1. Therefore,

$$T = \begin{pmatrix} t_{11} & \mathbf{0}^T \\ \mathbf{0} & D_1 \end{pmatrix}$$

a diagonal matrix.

Now here is a proof which doesn't involve block multiplication. Since T is normal, $T^*T = TT^*$. Writing this in terms of components and using the description of the adjoint as the transpose of the conjugate, yields the following for the ik^{th} entry of $T^*T = TT^*$.

$$\overbrace{\sum_j t_{ij} t_{jk}^* = \sum_j t_{ij} \overline{t_{kj}}}^{TT^*} = \overbrace{\sum_j t_{ij}^* t_{jk} = \sum_j \overline{t_{ji}} t_{jk}}^{T^*T}.$$

Now use the fact that T is upper triangular and let $i = k = 1$ to obtain the following from the above.

$$\sum_j |t_{1j}|^2 = \sum_j |t_{j1}|^2 = |t_{11}|^2$$

You see, $t_{j1} = 0$ unless $j = 1$ due to the assumption that T is upper triangular. This shows T is of the form

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}.$$

Now do the same thing only this time take $i = k = 2$ and use the result just established. Thus, from the above,

$$\sum_j |t_{2j}|^2 = \sum_j |t_{j2}|^2 = |t_{22}|^2,$$

showing that $t_{2j} = 0$ if $j > 2$ which means T has the form

$$\begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ 0 & * & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & * \end{pmatrix}.$$

Next let $i = k = 3$ and obtain that T looks like a diagonal matrix in so far as the first 3 rows and columns are concerned. Continuing in this way, it follows T is a diagonal matrix. ■

Theorem 7.4.11 *Let A be a normal matrix. Then there exists a unitary matrix U such that U^*AU is a diagonal matrix.*

Proof: From Theorem 7.4.4 there exists a unitary matrix U such that U^*AU equals an upper triangular matrix. The theorem is now proved if it is shown that the property of being normal is preserved under unitary similarity transformations. That is, verify that if A is normal and if $B = U^*AU$, then B is also normal. But this is easy.

$$\begin{aligned} B^*B &= U^*A^*UU^*AU = U^*A^*AU \\ &= U^*AA^*U = U^*AUU^*A^*U = BB^*. \end{aligned}$$

Therefore, U^*AU is a normal and upper triangular matrix and by Lemma 7.4.10 it must be a diagonal matrix. ■ The converse is also true. See Problem 9 below.

Corollary 7.4.12 *If A is Hermitian, then all the eigenvalues of A are real and there exists an orthonormal basis of eigenvectors.*

Proof: Since A is normal, there exists unitary, U such that $U^*AU = D$, a diagonal matrix whose diagonal entries are the eigenvalues of A. Therefore, $D^* = U^*A^*U = U^*AU = D$ showing D is real.

Finally, let

$$U = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{pmatrix}$$

where the $\mathbf{u}_i$ denote the columns of U and

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

The equation, $U^*AU = D$ implies

$$\begin{aligned} AU &= \begin{pmatrix} A\mathbf{u}_1 & A\mathbf{u}_2 & \cdots & A\mathbf{u}_n \end{pmatrix} \\ &= UD = \begin{pmatrix} \lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & \cdots & \lambda_n\mathbf{u}_n \end{pmatrix} \end{aligned}$$

where the entries denote the columns of AU and UD respectively. Therefore, $A\mathbf{u}_i = \lambda_i\mathbf{u}_i$ and since the matrix is unitary, the ij^{th} entry of U^*U equals δ_{ij} and so

$$\delta_{ij} = \mathbf{u}_i^*\mathbf{u}_j \equiv \mathbf{u}_j \cdot \mathbf{u}_i.$$

This proves the corollary because it shows the vectors $\{\mathbf{u}_i\}$ are orthonormal. Therefore, they form a basis because every orthonormal set of vectors is linearly independent. ■

Corollary 7.4.13 *If A is a real symmetric matrix, then A is Hermitian and there exists a real unitary matrix U such that $U^TAU = D$ where D is a diagonal matrix whose diagonal entries are the eigenvalues of A. By arranging the columns of U the diagonal entries of D can be made to appear in any order.*

Proof: This follows from Theorem 7.4.6 and Corollary 7.4.12. Let

$$U = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}$$

Then $AU = UD$ so

$$AU = \begin{pmatrix} A\mathbf{u}_1 & \cdots & A\mathbf{u}_n \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} D = \begin{pmatrix} \lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n \end{pmatrix}$$

Hence each column of U is an eigenvector of A. It follows that by rearranging these columns, the entries of D on the main diagonal can be made to appear in any order. To see this, consider such a rearrangement resulting in an orthogonal matrix U' given by

$$U' = \begin{pmatrix} \mathbf{u}_{i_1} & \cdots & \mathbf{u}_{i_n} \end{pmatrix}$$

Then

$$U'^TAU' = U'^T \begin{pmatrix} A\mathbf{u}_{i_1} & \cdots & A\mathbf{u}_{i_n} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{u}_{i_1}^T \\ \vdots \\ \mathbf{u}_{i_n}^T \end{pmatrix} \begin{pmatrix} \lambda_{i_1}\mathbf{u}_{i_1} & \cdots & \lambda_{i_n}\mathbf{u}_{i_n} \end{pmatrix} = \begin{pmatrix} \lambda_{i_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_{i_n} \end{pmatrix} ■$$

7.5 Trace And Determinant

The determinant has already been discussed. It is also clear that if $A = S^{-1}BS$ so that A, B are similar, then

$$\begin{aligned} \det(A) &= \det\left(S^{-1}\right)\det(S)\det(B) = \det\left(S^{-1}S\right)\det(B) \\ &= \det(I)\det(B) = \det(B) \end{aligned}$$

The **trace** is defined in the following definition.

Definition 7.5.1 *Let A be an $n \times n$ matrix whose ij^{th} entry is denoted as a_{ij}. Then*

$$\operatorname{trace}(A) \equiv \sum_i a_{ii}$$

In other words it is the sum of the entries down the main diagonal.

Theorem 7.5.2 *Let A be an $m \times n$ matrix and let B be an $n \times m$ matrix. Then*

$$\operatorname{trace}(AB) = \operatorname{trace}(BA).$$

Also if $B = S^{-1}AS$ so that A, B are similar, then

$$\operatorname{trace}(A) = \operatorname{trace}(B).$$

Proof:

$$\operatorname{trace}(AB) \equiv \sum_i \left(\sum_k A_{ik}B_{ki}\right) = \sum_k \sum_i B_{ki}A_{ik} = \operatorname{trace}(BA)$$

Therefore,

$$\operatorname{trace}(B) = \operatorname{trace}\left(S^{-1}AS\right) = \operatorname{trace}\left(ASS^{-1}\right) = \operatorname{trace}(A). \blacksquare$$

Theorem 7.5.3 *Let A be an $n \times n$ matrix. Then* $\operatorname{trace}(A)$ *equals the sum of the eigenvalues of A and* $\det(A)$ *equals the product of the eigenvalues of A.*

This is proved using Schur's theorem and is in Problem 17 below. Another important property of the trace is in the following theorem.

7.6 Quadratic Forms

Definition 7.6.1 *A quadratic form in three dimensions is an expression of the form*

$$\begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{7.13}$$

where A is a 3×3 symmetric matrix. In higher dimensions the idea is the same except you use a larger symmetric matrix in place of A. In two dimensions A is a 2×2 matrix.

For example, consider

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 3 & -4 & 1 \\ -4 & 0 & -4 \\ 1 & -4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{7.14}$$

which equals $3x^2 - 8xy + 2xz - 8yz + 3z^2$. This is very awkward because of the mixed terms such as $-8xy$. The idea is to pick different axes such that if x, y, z are taken with respect to these axes, the quadratic form is much simpler. In other words, look for new variables, x', y', and z' and a unitary matrix U such that

$$U \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{7.15}$$

and if you write the quadratic form in terms of the primed variables, there will be no mixed terms. Any symmetric real matrix is Hermitian and is therefore normal. From Corollary 7.4.13, it follows there exists a real unitary matrix U, (an orthogonal matrix) such that $U^T AU = D$ a diagonal matrix. Thus in the quadratic form, 7.13

$$\begin{aligned} \begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} x' & y' & z' \end{pmatrix} U^T AU \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \\ &= \begin{pmatrix} x' & y' & z' \end{pmatrix} D \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \end{aligned}$$

and in terms of these new variables, the quadratic form becomes

$$\lambda_1 (x')^2 + \lambda_2 (y')^2 + \lambda_3 (z')^2$$

where $D = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$. Similar considerations apply equally well in any other dimension. For the given example,

$$\begin{pmatrix} -\frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} & \frac{1}{3}\sqrt{6} & \frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{3} & -\frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \end{pmatrix} \begin{pmatrix} 3 & -4 & 1 \\ -4 & 0 & -4 \\ 1 & -4 & 3 \end{pmatrix} \cdot$$

$$\begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

and so if the new variables are given by

$$\begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

it follows that in terms of the new variables the quadratic form is $2(x')^2 - 4(y')^2 + 8(z')^2$. You can work other examples the same way.

7.7 Second Derivative Test

Under certain conditions the **mixed partial derivatives** will always be equal. This astonishing fact was first observed by Euler around 1734. It is also called Clairaut's theorem.

Theorem 7.7.1 *Suppose $f : U \subseteq \mathbb{F}^2 \to \mathbb{R}$ where U is an open set on which f_x, f_y, f_{xy} and f_{yx} exist. Then if f_{xy} and f_{yx} are continuous at the point $(x, y) \in U$, it follows*

$$f_{xy}(x, y) = f_{yx}(x, y).$$

Proof: Since U is open, there exists $r > 0$ such that $B((x, y), r) \subseteq U$. Now let $|t|, |s| < r/2$, t, s real numbers and consider

$$\Delta(s, t) \equiv \frac{1}{st}\{\overbrace{f(x+t, y+s) - f(x+t, y)}^{h(t)} - \overbrace{(f(x, y+s) - f(x, y))}^{h(0)}\}. \tag{7.16}$$

Note that $(x + t, y + s) \in U$ because

$$\begin{aligned} |(x+t, y+s) - (x, y)| &= |(t, s)| = \left(t^2 + s^2\right)^{1/2} \\ &\leq \left(\frac{r^2}{4} + \frac{r^2}{4}\right)^{1/2} = \frac{r}{\sqrt{2}} < r. \end{aligned}$$

As implied above, $h(t) \equiv f(x + t, y + s) - f(x + t, y)$. Therefore, by the mean value theorem from calculus and the (one variable) chain rule,

$$\begin{aligned} \Delta(s, t) &= \frac{1}{st}(h(t) - h(0)) = \frac{1}{st}h'(\alpha t)\, t \\ &= \frac{1}{s}\left(f_x(x + \alpha t, y + s) - f_x(x + \alpha t, y)\right) \end{aligned}$$

for some $\alpha \in (0, 1)$. Applying the mean value theorem again,

$$\Delta(s, t) = f_{xy}(x + \alpha t, y + \beta s)$$

where $\alpha, \beta \in (0, 1)$.

If the terms $f(x + t, y)$ and $f(x, y + s)$ are interchanged in 7.16, $\Delta(s, t)$ is unchanged and the above argument shows there exist $\gamma, \delta \in (0, 1)$ such that

$$\Delta(s, t) = f_{yx}(x + \gamma t, y + \delta s).$$

Letting $(s, t) \to (0, 0)$ and using the continuity of f_{xy} and f_{yx} at (x, y),

$$\lim_{(s,t)\to(0,0)} \Delta(s, t) = f_{xy}(x, y) = f_{yx}(x, y). \blacksquare$$

The following is obtained from the above by simply fixing all the variables except for the two of interest.

Corollary 7.7.2 *Suppose U is an open subset of $\mathbb{F}^n$ and $f : U \to \mathbb{R}$ has the property that for two indices, k, l, f_{x_k}, $f_{x_l}, f_{x_l x_k}$, and $f_{x_k x_l}$ exist on U and $f_{x_k x_l}$ and $f_{x_l x_k}$ are both continuous at $\mathbf{x} \in U$. Then $f_{x_k x_l}(\mathbf{x}) = f_{x_l x_k}(\mathbf{x})$.*

Thus the theorem asserts that the mixed partial derivatives are equal at $\mathbf{x}$ if they are defined near $\mathbf{x}$ and continuous at $\mathbf{x}$.

Now recall the Taylor formula with the Lagrange form of the remainder. What follows is a proof of this important result based on the mean value theorem or Rolle's theorem.

Theorem 7.7.3 *Suppose f has $n+1$ derivatives on an interval, (a, b) and let $c \in (a, b)$. Then if $x \in (a, b)$, there exists ξ between c and x such that*

$$f(x) = f(c) + \sum_{k=1}^{n} \frac{f^{(k)}(c)}{k!} (x-c)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-c)^{n+1}.$$

(In this formula, the symbol $\sum_{k=1}^{0} a_k$ will denote the number 0.)

Proof: If $n = 0$ then the theorem is true because it is just the mean value theorem. Suppose the theorem is true for $n-1, n \geq 1$. It can be assumed $x \neq c$ because if $x = c$ there is nothing to show. Then there exists K such that

$$f(x) - \left(f(c) + \sum_{k=1}^{n} \frac{f^{(k)}(c)}{k!} (x-c)^k + K (x-c)^{n+1} \right) = 0 \tag{7.17}$$

In fact,

$$K=\frac{-f\left(x\right)+\left(f\left(c\right)+\sum_{k=1}^{n}\frac{f^{\left(k\right)}\left(c\right)}{k!}\left(x-c\right)^{k}\right)}{\left(x-c\right)^{n+1}}.$$

Now define $F\left(t\right)$ for t in the closed interval determined by x and c by

$$F\left(t\right)\equiv f\left(x\right)-\left(f\left(t\right)+\sum_{k=1}^{n}\frac{f^{\left(k\right)}\left(c\right)}{k!}\left(x-t\right)^{k}+K\left(x-t\right)^{n+1}\right).$$

The c in 7.17 got replaced by t.

Therefore, $F(c) = 0$ by the way K was chosen and also $F(x) = 0$. By the mean value theorem or Rolle's theorem, there exists t_1 between x and c such that $F'(t_1) = 0$. Therefore,

$$\begin{aligned} 0 &= f'(t_1) - \sum_{k=1}^{n} \frac{f^{(k)}(c)}{k!} k (x-t_1)^{k-1} - K(n+1)(x-t_1)^n \\ &= f'(t_1) - \left(f'(c) + \sum_{k=1}^{n-1} \frac{f^{(k+1)}(c)}{k!} (x-t_1)^k \right) - K(n+1)(x-t_1)^n \\ &= f'(t_1) - \left(f'(c) + \sum_{k=1}^{n-1} \frac{f'^{(k)}(c)}{k!} (x-t_1)^k \right) - K(n+1)(x-t_1)^n \end{aligned}$$

By induction applied to f', there exists ξ between x and t_1 such that the above simplifies to

$$\begin{aligned} 0 &= \frac{f'^{(n)}(\xi)(x-t_1)^n}{n!} - K(n+1)(x-t_1)^n \\ &= \frac{f^{(n+1)}(\xi)(x-t_1)^n}{n!} - K(n+1)(x-t_1)^n \end{aligned}$$

therefore,

$$K = \frac{f^{(n+1)}(\xi)}{(n+1)\,n!} = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

and the formula is true for n. ■

The following is a special case and is what will be used.

Theorem 7.7.4 *Let $h : (-\delta, 1+\delta) \to \mathbb{R}$ have $m+1$ derivatives. Then there exists $t \in [0,1]$ such that*

$$h(1) = h(0) + \sum_{k=1}^{m} \frac{h^{(k)}(0)}{k!} + \frac{h^{(m+1)}(t)}{(m+1)!}.$$

Now let $f : U \to \mathbb{R}$ where $U \subseteq \mathbb{R}^n$ and suppose $f \in C^m(U)$. Let $\mathbf{x} \in U$ and let $r > 0$ be such that

$$B(\mathbf{x}, r) \subseteq U.$$

Then for $||\mathbf{v}|| < r$, consider

$$f(\mathbf{x}+t\mathbf{v}) - f(\mathbf{x}) \equiv h(t)$$

for $t \in [0,1]$. Then by the chain rule,

$$h'(t) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{x}+t\mathbf{v}) v_k, \;\; h''(t) = \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}+t\mathbf{v}) v_k v_j \;\; ■$$

Then from the Taylor formula stopping at the second derivative, the following theorem can be obtained.

Theorem 7.7.5 *Let $f : U \to \mathbb{R}$ and let $f \in C^2(U)$. Then if*

$$B(\mathbf{x}, r) \subseteq U,$$

and $||\mathbf{v}|| < r$, there exists $t \in (0,1)$ such that.

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{x}) v_k + \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}+t\mathbf{v}) v_k v_j \tag{7.18}$$

Definition 7.7.6 *Define the following matrix.*

$$H_{ij}(\mathbf{x}+t\mathbf{v}) \equiv \frac{\partial^2 f(\mathbf{x}+t\mathbf{v})}{\partial x_j \partial x_i}.$$

It is called the Hessian matrix. From Corollary 7.7.2, this is a symmetric matrix. Then in terms of this matrix, 7.18 can be written as

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x}) v_k + \frac{1}{2}\mathbf{v}^T H(\mathbf{x}+t\mathbf{v})\mathbf{v}$$

Then this implies $f(\mathbf{x}+\mathbf{v}) =$

$$f(\mathbf{x}) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x}) v_k + \frac{1}{2}\mathbf{v}^T H(\mathbf{x})\mathbf{v} + \frac{1}{2}\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right). \tag{7.19}$$

Using the above formula, here is the second derivative test.

Theorem 7.7.7 *In the above situation, suppose $f_{x_j}(\mathbf{x}) = 0$ for each x_j. Then if $H(\mathbf{x})$ has all positive eigenvalues, $\mathbf{x}$ is a local minimum for f. If $H(\mathbf{x})$ has all negative eigenvalues, then $\mathbf{x}$ is a local maximum. If $H(\mathbf{x})$ has a positive eigenvalue, then there exists a direction in which f has a local minimum at $\mathbf{x}$, while if $H(\mathbf{x})$ has a negative eigenvalue, there exists a direction in which $H(\mathbf{x})$ has a local maximum at $\mathbf{x}$.*

Proof: Since $f_{x_j}(\mathbf{x}) = 0$ for each x_j, formula 7.19 implies

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \frac{1}{2}\mathbf{v}^T H(\mathbf{x})\mathbf{v} + \frac{1}{2}\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right)$$

where $H(\mathbf{x})$ is a symmetric matrix. Thus, by Corollary 7.4.12 $H(\mathbf{x})$ has all real eigenvalues. Suppose first that $H(\mathbf{x})$ has all positive eigenvalues and that all are larger than $\delta^2 > 0$. Then $H(\mathbf{x})$ has an orthonormal basis of eigenvectors, $\{\mathbf{v}_i\}_{i=1}^n$ and if $\mathbf{u}$ is an arbitrary vector, $\mathbf{u} = \sum_{j=1}^n u_j \mathbf{v}_j$ where $u_j = \mathbf{u}\cdot\mathbf{v}_j$. Thus

$$\mathbf{u}^T H(\mathbf{x})\mathbf{u} = \left(\sum_{k=1}^n u_k \mathbf{v}_k^T\right) H(\mathbf{x}) \left(\sum_{j=1}^n u_j \mathbf{v}_j\right) = \sum_{j=1}^n u_j^2 \lambda_j \geq \delta^2 \sum_{j=1}^n u_j^2 = \delta^2 |\mathbf{u}|^2.$$

From 7.19 and the continuity of H, if $\mathbf{v}$ is small enough,

$$f(\mathbf{x}+\mathbf{v}) \geq f(\mathbf{x}) + \frac{1}{2}\delta^2 |\mathbf{v}|^2 - \frac{1}{4}\delta^2 |\mathbf{v}|^2 = f(\mathbf{x}) + \frac{\delta^2}{4}|\mathbf{v}|^2.$$

This shows the first claim of the theorem. The second claim follows from similar reasoning. Suppose $H(\mathbf{x})$ has a positive eigenvalue λ^2. Then let $\mathbf{v}$ be an eigenvector for this eigenvalue. From 7.19,

$$f(\mathbf{x}+t\mathbf{v}) = f(\mathbf{x}) + \frac{1}{2}t^2 \mathbf{v}^T H(\mathbf{x})\mathbf{v} + \frac{1}{2}t^2\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right)$$

which implies

$$\begin{aligned} f(\mathbf{x}+t\mathbf{v}) &= f(\mathbf{x}) + \frac{1}{2}t^2\lambda^2 |\mathbf{v}|^2 + \frac{1}{2}t^2\left(\mathbf{v}^T (H(\mathbf{x}+t\mathbf{v}) - H(\mathbf{x}))\mathbf{v}\right) \\ &\geq f(\mathbf{x}) + \frac{1}{4}t^2\lambda^2 |\mathbf{v}|^2 \end{aligned}$$

whenever t is small enough. Thus in the direction $\mathbf{v}$ the function has a local minimum at $\mathbf{x}$. The assertion about the local maximum in some direction follows similarly. ■

This theorem is an analogue of the second derivative test for higher dimensions. As in one dimension, when there is a zero eigenvalue, it may be impossible to determine from the Hessian matrix what the local qualitative behavior of the function is. For example, consider

$$f_1(x,y) = x^4 + y^2, \ f_2(x,y) = -x^4 + y^2.$$

Then $Df_i(0,0) = \mathbf{0}$ and for both functions, the Hessian matrix evaluated at $(0,0)$ equals

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

but the behavior of the two functions is very different near the origin. The second has a saddle point while the first has a minimum there.

7.8 The Estimation Of Eigenvalues

There are ways to estimate the eigenvalues for matrices. The most famous is known as Gerschgorin's theorem. This theorem gives a rough idea where the eigenvalues are just from looking at the matrix.

Theorem 7.8.1 *Let A be an $n \times n$ matrix. Consider the n Gerschgorin discs defined as*

$$D_i \equiv \left\{ \lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}.$$

Then every eigenvalue is contained in some Gerschgorin disc.

This theorem says to add up the absolute values of the entries of the i^{th} row which are off the main diagonal and form the disc centered at a_{ii} having this radius. The union of these discs contains $\sigma(A)$.

Proof: Suppose $A\mathbf{x} = \lambda \mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$. Then for $A = (a_{ij})$

$$\sum_{j \neq i} a_{ij} x_j = (\lambda - a_{ii}) x_i.$$

Therefore, picking k such that $|x_k| \geq |x_j|$ for all x_j, it follows that $|x_k| \neq 0$ since $|\mathbf{x}| \neq 0$ and

$$|x_k| \sum_{j \neq i} |a_{ij}| \geq \sum_{j \neq i} |a_{ij}| |x_j| \geq |\lambda - a_{ii}| |x_k| .$$

Now dividing by $|x_k|$, it follows λ is contained in the k^{th} Gerschgorin disc. ■

Example 7.8.2 *Here is a matrix. Estimate its eigenvalues.*

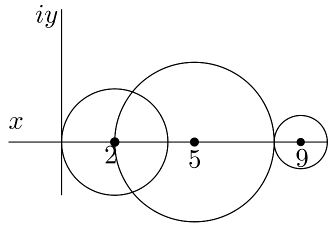
$$\begin{pmatrix} 2 & 1 & 1 \\ 3 & 5 & 0 \\ 0 & 1 & 9 \end{pmatrix}$$

According to Gerschgorin's theorem the eigenvalues are contained in the disks

$$D_1 = \{\lambda \in \mathbb{C} : |\lambda - 2| \leq 2\}, D_2 = \{\lambda \in \mathbb{C} : |\lambda - 5| \leq 3\},$$

$$D_3 = \{\lambda \in \mathbb{C} : |\lambda - 9| \leq 1\}$$

It is important to observe that these disks are in the complex plane. In general this is the case. If you want to find eigenvalues they will be complex numbers.



So what are the values of the eigenvalues? In this case they are real. You can compute them by graphing the characteristic polynomial, $\lambda^3 - 16\lambda^2 + 70\lambda - 66$ and then zooming in on the zeros. If you do this you find the solution is $\{\lambda = 1.2953\}, \{\lambda = 5.5905\}$, $\{\lambda = 9.1142\}$. Of course these are only approximations and so this information is useless for finding eigenvectors. However, in many applications, it is the size of the eigenvalues which is important and so these numerical values would be helpful for such applications. In this case, you might think there is no real reason for Gerschgorin's theorem. Why not just compute the characteristic equation and graph and zoom? This is fine up to a point, but what if the matrix was huge? Then it might be hard to find the characteristic polynomial. Remember the difficulties in expanding a big matrix along a row or column. Also, what if the eigenvalues were complex? You don't see these by following this procedure. However, Gerschgorin's theorem will at least estimate them.

7.9 Advanced Theorems

More can be said but this requires some theory from complex variables[1]. The following is a fundamental theorem about counting zeros.

Theorem 7.9.1 *Let U be a region and let $\gamma : [a, b] \to U$ be closed, continuous, bounded variation, and the winding number, $n(\gamma, z) = 0$ for all $z \notin U$. Suppose also that f is*

[1] If you haven't studied the theory of a complex variable, you should skip this section because you won't understand any of it.

analytic on U having zeros $a_1, \cdots, a_m$ where the zeros are repeated according to multiplicity, and suppose that none of these zeros are on $\gamma([a,b])$. Then

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}dz = \sum_{k=1}^m n(\gamma, a_k).$$

Proof: It is given that $f(z) = \prod_{j=1}^m (z - a_j) g(z)$ where $g(z) \neq 0$ on U. Hence using the product rule,

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^m \frac{1}{z - a_j} + \frac{g'(z)}{g(z)}$$

where $\frac{g'(z)}{g(z)}$ is analytic on U and so

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}dz = \sum_{j=1}^m n(\gamma, a_j) + \frac{1}{2\pi i}\int_\gamma \frac{g'(z)}{g(z)}dz = \sum_{j=1}^m n(\gamma, a_j). \blacksquare$$

Now let A be an $n \times n$ matrix. Recall that the eigenvalues of A are given by the zeros of the polynomial, $p_A(z) = \det(zI - A)$ where I is the $n \times n$ identity. You can argue that small changes in A will produce small changes in $p_A(z)$ and $p'_A(z)$. Let γ_k denote a very small closed circle which winds around z_k, one of the eigenvalues of A, in the counter clockwise direction so that $n(\gamma_k, z_k) = 1$. This circle is to enclose only z_k and is to have no other eigenvalue on it. Then apply Theorem 7.9.1. According to this theorem

$$\frac{1}{2\pi i}\int_\gamma \frac{p'_A(z)}{p_A(z)}dz$$

is always an integer equal to the multiplicity of z_k as a root of $p_A(t)$. Therefore, small changes in A result in no change to the above contour integral because it must be an integer and small changes in A result in small changes in the integral. Therefore whenever B is close enough to A, the two matrices have the same number of zeros inside γ_k, the zeros being counted according to multiplicity. By making the radius of the small circle equal to ε where ε is less than the minimum distance between any two distinct eigenvalues of A, this shows that if B is close enough to A, every eigenvalue of B is closer than ε to some eigenvalue of A. $\blacksquare$

Theorem 7.9.2 *If λ is an eigenvalue of A, then if all the entries of B are close enough to the corresponding entries of A, some eigenvalue of B will be within ε of λ.*

Consider the situation that $A(t)$ is an $n \times n$ matrix and that $t \to A(t)$ is continuous for $t \in [0,1]$.

Lemma 7.9.3 *Let $\lambda(t) \in \sigma(A(t))$ for $t < 1$ and let $\Sigma_t = \cup_{s \geq t} \sigma(A(s))$. Also let K_t be the connected component of $\lambda(t)$ in Σ_t. Then there exists $\eta > 0$ such that $K_t \cap \sigma(A(s)) \neq \emptyset$ for all $s \in [t, t+\eta]$.*

Proof: Denote by $D(\lambda(t), \delta)$ the disc centered at $\lambda(t)$ having radius $\delta > 0$, with other occurrences of this notation being defined similarly. Thus

$$D(\lambda(t), \delta) \equiv \{z \in \mathbb{C} : |\lambda(t) - z| \leq \delta\}.$$

Suppose $\delta > 0$ is small enough that $\lambda(t)$ is the only element of $\sigma(A(t))$ contained in $D(\lambda(t), \delta)$ and that $p_{A(t)}$ has no zeroes on the boundary of this disc. Then by continuity, and the above discussion and theorem, there exists $\eta > 0, t + \eta < 1$, such that for $s \in [t, t+\eta]$, $p_{A(s)}$ also has no zeroes on the boundary of this disc and $A(s)$ has the same number of eigenvalues, counted according to multiplicity, in the disc as $A(t)$. Thus $\sigma(A(s)) \cap D(\lambda(t), \delta) \neq \emptyset$ for all $s \in [t, t+\eta]$. Now let

$$H = \bigcup_{s \in [t, t+\eta]} \sigma(A(s)) \cap D(\lambda(t), \delta).$$

It will be shown that H is connected. Suppose not. Then $H = P \cup Q$ where P, Q are separated and $\lambda(t) \in P$. Let $s_0 \equiv \inf\{s : \lambda(s) \in Q \text{ for some } \lambda(s) \in \sigma(A(s))\}$. There exists $\lambda(s_0) \in \sigma(A(s_0)) \cap D(\lambda(t), \delta)$. If $\lambda(s_0) \notin Q$, then from the above discussion there are $\lambda(s) \in \sigma(A(s)) \cap Q$ for $s > s_0$ arbitrarily close to $\lambda(s_0)$. Therefore, $\lambda(s_0) \in Q$ which shows that $s_0 > t$ because $\lambda(t)$ is the only element of $\sigma(A(t))$ in $D(\lambda(t), \delta)$ and $\lambda(t) \in P$. Now let $s_n \uparrow s_0$. Then $\lambda(s_n) \in P$ for any $\lambda(s_n) \in \sigma(A(s_n)) \cap D(\lambda(t), \delta)$ and also it follows from the above discussion that for some choice of $s_n \to s_0$, $\lambda(s_n) \to \lambda(s_0)$ which contradicts P and Q separated and nonempty. Since P is nonempty, this shows $Q = \emptyset$. Therefore, H is connected as claimed. But $K_t \supseteq H$ and so $K_t \cap \sigma(A(s)) \neq \emptyset$ for all $s \in [t, t+\eta]$. ■

Theorem 7.9.4 *Suppose $A(t)$ is an $n \times n$ matrix and that $t \to A(t)$ is continuous for $t \in [0,1]$. Let $\lambda(0) \in \sigma(A(0))$ and define $\Sigma \equiv \cup_{t \in [0,1]} \sigma(A(t))$. Let $K_{\lambda(0)} = K_0$ denote the connected component of $\lambda(0)$ in Σ. Then $K_0 \cap \sigma(A(t)) \neq \emptyset$ for all $t \in [0,1]$.*

Proof: Let $S \equiv \{t \in [0,1] : K_0 \cap \sigma(A(s)) \neq \emptyset \text{ for all } s \in [0,t]\}$. Then $0 \in S$. Let $t_0 = \sup(S)$. Say $\sigma(A(t_0)) = \lambda_1(t_0), \cdots, \lambda_r(t_0)$.

Claim: At least one of these is a limit point of K_0 and consequently must be in K_0 which shows that S has a last point. Why is this claim true? Let $s_n \uparrow t_0$ so $s_n \in S$. Now let the discs, $D(\lambda_i(t_0), \delta), i = 1, \cdots, r$ be disjoint with $p_{A(t_0)}$ having no zeroes on γ_i the boundary of $D(\lambda_i(t_0), \delta)$. Then for n large enough it follows from Theorem 7.9.1 and the discussion following it that $\sigma(A(s_n))$ is contained in $\cup_{i=1}^r D(\lambda_i(t_0), \delta)$. It follows that $K_0 \cap (\sigma(A(t_0)) + D(0, \delta)) \neq \emptyset$ for all δ small enough. This requires at least one of the $\lambda_i(t_0)$ to be in $\overline{K_0}$. Therefore, $t_0 \in S$ and S has a last point.

Now by Lemma 7.9.3, if $t_0 < 1$, then $K_0 \cup K_t$ would be a strictly larger connected set containing $\lambda(0)$. (The reason this would be strictly larger is that $K_0 \cap \sigma(A(s)) = \emptyset$ for some $s \in (t, t+\eta)$ while $K_t \cap \sigma(A(s)) \neq \emptyset$ for all $s \in [t, t+\eta]$.) Therefore, $t_0 = 1$. ■

Corollary 7.9.5 *Suppose one of the Gerschgorin discs, D_i is disjoint from the union of the others. Then D_i contains an eigenvalue of A. Also, if there are n disjoint Gerschgorin discs, then each one contains an eigenvalue of A.*

Proof: Denote by $A(t)$ the matrix $\left(a_{ij}^t\right)$ where if $i \neq j$, $a_{ij}^t = ta_{ij}$ and $a_{ii}^t = a_{ii}$. Thus to get $A(t)$ multiply all non diagonal terms by t. Let $t \in [0,1]$. Then $A(0) = \text{diag}(a_{11}, \cdots, a_{nn})$ and $A(1) = A$. Furthermore, the map, $t \to A(t)$ is continuous. Denote by D_j^t the Gerschgorin disc obtained from the j^{th} row for the matrix $A(t)$. Then it is clear that $D_j^t \subseteq D_j$ the j^{th} Gerschgorin disc for A. It follows a_{ii} is the eigenvalue for $A(0)$ which is contained in the disc, consisting of the single point a_{ii} which is contained in D_i. Letting K be the connected component in Σ for Σ defined in Theorem 7.9.4 which is determined by a_{ii}, Gerschgorin's theorem implies that $K \cap \sigma(A(t)) \subseteq \cup_{j=1}^n D_j^t \subseteq \cup_{j=1}^n D_j = D_i \cup (\cup_{j \neq i} D_j)$ and also, since K is connected, there are not points of K in both D_i and $(\cup_{j \neq i} D_j)$. Since at least one point of K is in D_i,(a_{ii}), it follows all of K must be contained in D_i. Now by Theorem 7.9.4 this shows there are points of $K \cap \sigma(A)$ in D_i. The last assertion follows immediately. ■

This can be improved even more. This involves the following lemma.

Lemma 7.9.6 *In the situation of Theorem 7.9.4 suppose* $\lambda(0) = K_0 \cap \sigma(A(0))$ *and that*

$\lambda(0)$ *is a simple root of the characteristic equation of* $A(0)$. *Then for all* $t \in [0,1]$,

$$\sigma(A(t)) \cap K_0 = \lambda(t)$$

where $\lambda(t)$ *is a simple root of the characteristic equation of* $A(t)$.

Proof: Let $S \equiv \{t \in [0,1] : K_0 \cap \sigma(A(s)) = \lambda(s)$, a simple eigenvalue for all $s \in [0,t]\}$. Then $0 \in S$ so it is nonempty. Let $t_0 = \sup(S)$ and suppose $\lambda_1 \neq \lambda_2$ are two elements of $\sigma(A(t_0)) \cap K_0$. Then choosing $\eta > 0$ small enough, and letting D_i be disjoint discs containing λ_i respectively, similar arguments to those of Lemma 7.9.3 can be used to conclude

$$H_i \equiv \cup_{s \in [t_0-\eta, t_0]} \sigma(A(s)) \cap D_i$$

is a connected and nonempty set for $i = 1,2$ which would require that $H_i \subseteq K_0$. But then there would be two different eigenvalues of $A(s)$ contained in K_0, contrary to the definition of t_0. Therefore, there is at most one eigenvalue $\lambda(t_0) \in K_0 \cap \sigma(A(t_0))$. Could it be a repeated root of the characteristic equation? Suppose $\lambda(t_0)$ is a repeated root of the characteristic equation. As before, choose a small disc, D centered at $\lambda(t_0)$ and η small enough that

$$H \equiv \cup_{s \in [t_0-\eta, t_0]} \sigma(A(s)) \cap D$$

is a nonempty connected set containing either multiple eigenvalues of $A(s)$ or else a single repeated root to the characteristic equation of $A(s)$. But since H is connected and contains $\lambda(t_0)$ it must be contained in K_0 which contradicts the condition for $s \in S$ for all these $s \in [t_0 - \eta, t_0]$. Therefore, $t_0 \in S$ as hoped. If $t_0 < 1$, there exists a small disc centered at $\lambda(t_0)$ and $\eta > 0$ such that for all $s \in [t_0, t_0 + \eta]$, $A(s)$ has only simple eigenvalues in D and the only eigenvalues of $A(s)$ which could be in K_0 are in D. (This last assertion follows from noting that $\lambda(t_0)$ is the only eigenvalue of $A(t_0)$ in K_0 and so the others are at a positive distance from K_0. For s close enough to t_0, the eigenvalues of $A(s)$ are either close to these eigenvalues of $A(t_0)$ at a positive distance from K_0 or they are close to the eigenvalue $\lambda(t_0)$ in which case it can be assumed they are in D.) But this shows that t_0 is not really an upper bound to S. Therefore, $t_0 = 1$ and the lemma is proved. ■

With this lemma, the conclusion of the above corollary can be sharpened.

Corollary 7.9.7 *Suppose one of the Gerschgorin discs,* D_i *is disjoint from the union of the others. Then* D_i *contains exactly one eigenvalue of* A *and this eigenvalue is a simple root to the characteristic polynomial of* A.

Proof: In the proof of Corollary 7.9.5, note that a_{ii} is a simple root of $A(0)$ since otherwise the i^{th} Gerschgorin disc would not be disjoint from the others. Also, K, the connected component determined by a_{ii} must be contained in D_i because it is connected and by Gerschgorin's theorem above, $K \cap \sigma(A(t))$ must be contained in the union of the Gerschgorin discs. Since all the other eigenvalues of $A(0)$, the a_{jj}, are outside D_i, it follows that $K \cap \sigma(A(0)) = a_{ii}$. Therefore, by Lemma 7.9.6, $K \cap \sigma(A(1)) = K \cap \sigma(A)$ consists of a single simple eigenvalue. ■

Example 7.9.8 *Consider the matrix*

$$\begin{pmatrix} 5 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

The Gerschgorin discs are $D(5,1), D(1,2)$, and $D(0,1)$. Observe $D(5,1)$ is disjoint from the other discs. Therefore, there should be an eigenvalue in $D(5,1)$. The actual eigenvalues are not easy to find. They are the roots of the characteristic equation, $t^3 - 6t^2 + 3t + 5 = 0$. The numerical values of these are $-.669\,66, 1.423\,1$, and $5.246\,55$, verifying the predictions of Gerschgorin's theorem.

7.10 Exercises

1. Explain why it is typically impossible to compute the upper triangular matrix whose existence is guaranteed by Schur's theorem.

2. Now recall the QR factorization of Theorem 5.7.5 on Page 174. The QR algorithm is a technique which does compute the upper triangular matrix in Schur's theorem. There is much more to the QR algorithm than will be presented here. In fact, what I am about to show you is not the way it is done in practice. One first obtains what is called a Hessenburg matrix for which the algorithm will work better. However, the idea is as follows. Start with A an $n \times n$ matrix having real eigenvalues. Form $A = QR$ where Q is orthogonal and R is upper triangular. (Right triangular.) This can be done using the technique of Theorem 5.7.5 using Householder matrices. Next take $A_1 \equiv RQ$. Show that $A = QA_1Q^T$. In other words these two matrices, A, A_1 are similar. Explain why they have the same eigenvalues. Continue by letting A_1 play the role of A. Thus the algorithm is of the form $A_n = QR_n$ and $A_{n+1} = R_{n+1}Q$. Explain why $A = Q_nA_nQ_n^T$ for some Q_n orthogonal. Thus A_n is a sequence of matrices each similar to A. The remarkable thing is that often these matrices converge to an upper triangular matrix T and $A = QTQ^T$ for some orthogonal matrix, the limit of the Q_n where the limit means the entries converge. Then the process computes the upper triangular Schur form of the matrix A. Thus the eigenvalues of A appear on the diagonal of T. You will see approximately what these are as the process continues.

3. ↑Try the QR algorithm on
$$\begin{pmatrix} -1 & -2 \\ 6 & 6 \end{pmatrix}$$
which has eigenvalues 3 and 2. I suggest you use a computer algebra system to do the computations.

4. ↑Now try the QR algorithm on
$$\begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix}$$
Show that the algorithm cannot converge for this example. **Hint:** Try a few iterations of the algorithm. Use a computer algebra system if you like.

5. ↑Show the two matrices $A \equiv \begin{pmatrix} 0 & -1 \\ 4 & 0 \end{pmatrix}$ and $B \equiv \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$ are similar; that is there exists a matrix S such that $A = S^{-1}BS$ but there is no orthogonal matrix Q such that $Q^TBQ = A$. Show the QR algorithm does converge for the matrix B although it fails to do so for A.

6. Let F be an $m \times n$ matrix. Show that F^*F has all real eigenvalues and furthermore, they are all nonnegative.

7. If A is a real $n \times n$ matrix and λ is a complex eigenvalue $\lambda = a + ib, b \neq 0$, of A having eigenvector $\mathbf{z} + i\mathbf{w}$, show that $\mathbf{w} \neq \mathbf{0}$.

8. Suppose $A = Q^TDQ$ where Q is an orthogonal matrix and all the matrices are real. Also D is a diagonal matrix. Show that A must be symmetric.

9. Suppose A is an $n \times n$ matrix and there exists a unitary matrix U such that
$$A = U^*DU$$
where D is a diagonal matrix. Explain why A must be normal.

10. If A is Hermitian, show that $\det(A)$ must be real.

11. Show that every unitary matrix preserves distance. That is, if U is unitary,
$$|U\mathbf{x}| = |\mathbf{x}|.$$

12. Show that if a matrix does preserve distances, then it must be unitary.

13. $\uparrow$Show that a complex normal matrix A is unitary if and only if its eigenvalues have magnitude equal to 1.

14. Suppose A is an $n \times n$ matrix which is diagonally dominant. Recall this means
$$\sum_{j \neq i} |a_{ij}| < |a_{ii}|$$
show A^{-1} must exist.

15. Give some disks in the complex plane whose union contains all the eigenvalues of the matrix
$$\begin{pmatrix} 1+2i & 4 & 2 \\ 0 & i & 3 \\ 5 & 6 & 7 \end{pmatrix}$$

16. Show a square matrix is invertible if and only if it has no zero eigenvalues.

17. Using Schur's theorem, show the trace of an $n \times n$ matrix equals the sum of the eigenvalues and the determinant of an $n \times n$ matrix is the product of the eigenvalues.

18. Using Schur's theorem, show that if A is any complex $n \times n$ matrix having eigenvalues $\{\lambda_i\}$ listed according to multiplicity, then $\sum_{i,j} |A_{ij}|^2 \geq \sum_{i=1}^n |\lambda_i|^2$. Show that equality holds if and only if A is normal. This inequality is called Schur's inequality. [19]

19. Here is a matrix.
$$\begin{pmatrix} 1234 & 6 & 5 & 3 \\ 0 & -654 & 9 & 123 \\ 98 & 123 & 10,000 & 11 \\ 56 & 78 & 98 & 400 \end{pmatrix}$$

I know this matrix has an inverse before doing any computations. How do I know?

20. Show the critical points of the following function are

$$(0,-3,0)\,,(2,-3,0)\,,\text{and }\left(1,-3,-\frac{1}{3}\right)$$

and classify them as local minima, local maxima or saddle points.
$f(x,y,z) = -\frac{3}{2}x^4 + 6x^3 - 6x^2 + zx^2 - 2zx - 2y^2 - 12y - 18 - \frac{3}{2}z^2.$

21. Here is a function of three variables.

$$f(x,y,z) = 13x^2 + 2xy + 8xz + 13y^2 + 8yz + 10z^2$$

change the variables so that in the new variables there are no mixed terms, terms involving xy, yz etc. Two eigenvalues are 12 and 18.

22. Here is a function of three variables.

$$f(x,y,z) = 2x^2 - 4x + 2 + 9yx - 9y - 3zx + 3z + 5y^2 - 9zy - 7z^2$$

change the variables so that in the new variables there are no mixed terms, terms involving xy, yz etc. The eigenvalues of the matrix which you will work with are $-\frac{17}{2}, \frac{19}{2}, -1$.

23. Here is a function of three variables.

$$f(x, y, z) = -x^2 + 2xy + 2xz - y^2 + 2yz - z^2 + x$$

change the variables so that in the new variables there are no mixed terms, terms involving xy, yz etc.

24. Show the critical points of the function,

$$f(x, y, z) = -2yx^2 - 6yx - 4zx^2 - 12zx + y^2 + 2yz.$$

are points of the form,

$$(x, y, z) = \left(t, 2t^2 + 6t, -t^2 - 3t\right)$$

for $t \in \mathbb{R}$ and classify them as local minima, local maxima or saddle points.

25. Show the critical points of the function

$$f(x, y, z) = \frac{1}{2}x^4 - 4x^3 + 8x^2 - 3zx^2 + 12zx + 2y^2 + 4y + 2 + \frac{1}{2}z^2.$$

are $(0, -1, 0), (4, -1, 0)$, and $(2, -1, -12)$ and classify them as local minima, local maxima or saddle points.

26. Let $f(x, y) = 3x^4 - 24x^2 + 48 - yx^2 + 4y$. Find and classify the critical points using the second derivative test.

27. Let $f(x, y) = 3x^4 - 5x^2 + 2 - y^2x^2 + y^2$. Find and classify the critical points using the second derivative test.

28. Let $f(x, y) = 5x^4 - 7x^2 - 2 - 3y^2x^2 + 11y^2 - 4y^4$. Find and classify the critical points using the second derivative test.

29. Let $f(x, y, z) = -2x^4 - 3yx^2 + 3x^2 + 5x^2z + 3y^2 - 6y + 3 - 3zy + 3z + z^2$. Find and classify the critical points using the second derivative test.

30. Let $f(x, y, z) = 3yx^2 - 3x^2 - x^2z - y^2 + 2y - 1 + 3zy - 3z - 3z^2$. Find and classify the critical points using the second derivative test.

31. Let Q be orthogonal. Find the possible values of $\det(Q)$.

32. Let U be unitary. Find the possible values of $\det(U)$.

33. If a matrix is nonzero can it have only zero for eigenvalues?

34. A matrix A is called nilpotent if $A^k = 0$ for some positive integer k. Suppose A is a nilpotent matrix. Show it has only 0 for an eigenvalue.

35. If A is a nonzero nilpotent matrix, show it must be defective.

36. Suppose A is a nondefective $n \times n$ matrix and its eigenvalues are all either 0 or 1. Show $A^2 = A$. Could you say anything interesting if the eigenvalues were all either 0,1,or -1? By DeMoivre's theorem, an n^{th} root of unity is of the form

$$\left(\cos\left(\frac{2k\pi}{n}\right) + i\sin\left(\frac{2k\pi}{n}\right)\right)$$

Could you generalize the sort of thing just described to get $A^n = A$? **Hint:** Since A is nondefective, there exists S such that $S^{-1}AS = D$ where D is a diagonal matrix.

37. This and the following problems will present most of a differential equations course. Most of the explanations are given. You fill in any details needed. To begin with, consider the scalar initial value problem

$$y' = ay,\ y(t_0) = y_0$$

When a is real, show the unique solution to this problem is $y = y_0 e^{a(t-t_0)}$. Next suppose

$$y' = (a+ib)\,y,\ y(t_0) = y_0 \tag{7.20}$$

where $y(t) = u(t) + iv(t)$. Show there exists a unique solution and it is given by $y(t) =$

$$y_0 e^{a(t-t_0)}\left(\cos b(t-t_0) + i\sin b(t-t_0)\right) \equiv e^{(a+ib)(t-t_0)}y_0. \tag{7.21}$$

Next show that for a real or complex there exists a unique solution to the initial value problem

$$y' = ay + f,\ y(t_0) = y_0$$

and it is given by

$$y(t) = e^{a(t-t_0)}y_0 + e^{at}\int_{t_0}^{t} e^{-as}f(s)\,ds.$$

Hint: For the first part write as $y' - ay = 0$ and multiply both sides by e^{-at}. Then explain why you get

$$\frac{d}{dt}\left(e^{-at}y(t)\right) = 0,\ y(t_0) = 0.$$

Now you finish the argument. To show uniqueness in the second part, suppose

$$y' = (a+ib)\,y,\ y(t_0) = 0$$

and verify this requires $y(t) = 0$. To do this, note

$$\overline{y}' = (a-ib)\,\overline{y},\ \overline{y}(t_0) = 0$$

and that $|y|^2(t_0) = 0$ and

$$\frac{d}{dt}|y(t)|^2 = y'(t)\,\overline{y}(t) + \overline{y}'(t)\,y(t)$$

$$= (a+ib)\,y(t)\,\overline{y}(t) + (a-ib)\,\overline{y}(t)\,y(t) = 2a\,|y(t)|^2.$$

Thus from the first part $|y(t)|^2 = 0e^{-2at} = 0$. Finally observe by a simple computation that 7.20 is solved by 7.21. For the last part, write the equation as

$$y' - ay = f$$

and multiply both sides by e^{-at} and then integrate from t_0 to t using the initial condition.

38. Now consider A an $n \times n$ matrix. By Schur's theorem there exists unitary Q such that

$$Q^{-1}AQ = T$$

where T is upper triangular. Now consider the first order initial value problem

$$\mathbf{x}' = A\mathbf{x},\ \mathbf{x}(t_0) = \mathbf{x}_0.$$

Show there exists a unique solution to this first order system. **Hint:** Let $\mathbf{y} = Q^{-1}\mathbf{x}$ and so the system becomes

$$\mathbf{y}' = T\mathbf{y},\ \mathbf{y}(t_0) = Q^{-1}\mathbf{x}_0 \tag{7.22}$$

Now letting $\mathbf{y} = (y_1, \cdots, y_n)^T$, the bottom equation becomes

$$y_n' = t_{nn}y_n,\ y_n(t_0) = \left(Q^{-1}\mathbf{x}_0\right)_n.$$

Then use the solution you get in this to get the solution to the initial value problem which occurs one level up, namely

$$y_{n-1}' = t_{(n-1)(n-1)}y_{n-1} + t_{(n-1)n}y_n,\ y_{n-1}(t_0) = \left(Q^{-1}\mathbf{x}_0\right)_{n-1}$$

Continue doing this to obtain a unique solution to 7.22.

39. Now suppose $\Phi(t)$ is an $n \times n$ matrix of the form

$$\Phi(t) = \begin{pmatrix} \mathbf{x}_1(t) & \cdots & \mathbf{x}_n(t) \end{pmatrix} \tag{7.23}$$

where

$$\mathbf{x}_k'(t) = A\mathbf{x}_k(t).$$

Explain why

$$\Phi'(t) = A\Phi(t)$$

if and only if $\Phi(t)$ is given in the form of 7.23. Also explain why if $\mathbf{c} \in \mathbb{F}^n$, $\mathbf{y}(t) \equiv \Phi(t)\mathbf{c}$

solves the equation $\mathbf{y}'(t) = A\mathbf{y}(t)$.

40. In the above problem, consider the question whether all solutions to

$$\mathbf{x}' = A\mathbf{x} \tag{7.24}$$

are obtained in the form $\Phi(t)\mathbf{c}$ for some choice of $\mathbf{c} \in \mathbb{F}^n$. In other words, is the general solution to this equation $\Phi(t)\mathbf{c}$ for $\mathbf{c} \in \mathbb{F}^n$? Prove the following theorem using linear algebra.

Theorem 7.10.1 *Suppose $\Phi(t)$ is an $n \times n$ matrix which satisfies $\Phi'(t) = A\Phi(t)$. Then the general solution to 7.24 is $\Phi(t)\mathbf{c}$ if and only if $\Phi(t)^{-1}$ exists for some t. Furthermore, if $\Phi'(t) = A\Phi(t)$, then either $\Phi(t)^{-1}$ exists for all t or $\Phi(t)^{-1}$ never exists for any t.*

($\det(\Phi(t))$ is called the Wronskian and this theorem is sometimes called the Wronskian alternative.)

Hint: Suppose first the general solution is of the form $\Phi(t)\mathbf{c}$ where $\mathbf{c}$ is an arbitrary constant vector in $\mathbb{F}^n$. You need to verify $\Phi(t)^{-1}$ exists for some t. In fact, show $\Phi(t)^{-1}$ exists for every t. Suppose then that $\Phi(t_0)^{-1}$ does not exist. Explain why there exists $\mathbf{c} \in \mathbb{F}^n$ such that there is no solution $\mathbf{x}$ to the equation $\mathbf{c} = \Phi(t_0)\mathbf{x}$. By the existence part of Problem 38 there exists a solution to

$$\mathbf{x}' = A\mathbf{x},\ \mathbf{x}(t_0) = \mathbf{c}$$

but this cannot be in the form $\Phi(t)\mathbf{c}$. Thus for every t, $\Phi(t)^{-1}$ exists. Next suppose for some $t_0, \Phi(t_0)^{-1}$ exists. Let $\mathbf{z}' = A\mathbf{z}$ and choose $\mathbf{c}$ such that

$$\mathbf{z}(t_0) = \Phi(t_0)\mathbf{c}$$

Then both $\mathbf{z}(t), \Phi(t)\mathbf{c}$ solve

$$\mathbf{x}' = A\mathbf{x},\ \mathbf{x}(t_0) = \mathbf{z}(t_0)$$

Apply uniqueness to conclude $\mathbf{z} = \Phi(t)\mathbf{c}$. Finally, consider that $\Phi(t)\mathbf{c}$ for $\mathbf{c} \in \mathbb{F}^n$ either is the general solution or it is not the general solution. If it is, then $\Phi(t)^{-1}$ exists for all t. If it is not, then $\Phi(t)^{-1}$ cannot exist for any t from what was just shown.

41. Let $\Phi'(t) = A\Phi(t)$. Then $\Phi(t)$ is called a fundamental matrix if $\Phi(t)^{-1}$ exists for all t. Show there exists a unique solution to the equation

$$\mathbf{x}' = A\mathbf{x} + \mathbf{f},\ \mathbf{x}(t_0) = \mathbf{x}_0 \tag{7.25}$$

and it is given by the formula

$$\mathbf{x}(t) = \Phi(t)\Phi(t_0)^{-1}\mathbf{x}_0 + \Phi(t)\int_{t_0}^{t}\Phi(s)^{-1}\mathbf{f}(s)\,ds$$

Now these few problems have done virtually everything of significance in an entire undergraduate differential equations course, illustrating the superiority of linear algebra. The above formula is called the variation of constants formula.

Hint: Uniquenss is easy. If $\mathbf{x}_1, \mathbf{x}_2$ are two solutions then let $\mathbf{u}(t) = \mathbf{x}_1(t) - \mathbf{x}_2(t)$ and argue $\mathbf{u}' = A\mathbf{u},\ \mathbf{u}(t_0) = \mathbf{0}$. Then use Problem 38. To verify there exists a solution, you could just differentiate the above formula using the fundamental theorem of calculus and verify it works. Another way is to assume the solution in the form

$$\mathbf{x}(t) = \Phi(t)\mathbf{c}(t)$$

and find $\mathbf{c}(t)$ to make it all work out. This is called the method of variation of parameters.

42. Show there exists a special Φ such that $\Phi'(t) = A\Phi(t),\ \Phi(0) = I$, and suppose $\Phi(t)^{-1}$ exists for all t. Show using uniqueness that

$$\Phi(-t) = \Phi(t)^{-1}$$

and that for all $t, s \in \mathbb{R}$

$$\Phi(t+s) = \Phi(t)\Phi(s)$$

Explain why with this special Φ, the solution to 7.25 can be written as

$$\mathbf{x}(t) = \Phi(t - t_0)\mathbf{x}_0 + \int_{t_0}^{t} \Phi(t - s)\mathbf{f}(s)\,ds.$$

Hint: Let $\Phi(t)$ be such that the j^{th} column is $\mathbf{x}_j(t)$ where

$$\mathbf{x}_j' = A\mathbf{x}_j,\ \mathbf{x}_j(0) = \mathbf{e}_j.$$

Use uniqueness as required.

43. You can see more on this problem and the next one in the latest version of Horn and Johnson, [16]. Two $n \times n$ matrices A, B are said to be congruent if there is an invertible P such that

$$B = PAP^*$$

Let A be a Hermitian matrix. Thus it has all real eigenvalues. Let n_+ be the number of positive eigenvalues, n_-, the number of negative eigenvalues and n_0 the number of zero eigenvalues. For k a positive integer, let I_k denote the $k \times k$ identity matrix and O_k the $k \times k$ zero matrix. Then the inertia matrix of A is the following block diagonal $n \times n$ matrix.

$$\begin{pmatrix} I_{n_+} & & \\ & I_{n_-} & \\ & & O_{n_0} \end{pmatrix}$$

Show that A is congruent to its inertia matrix. Next show that congruence is an equivalence relation on the set of Hermitian matrices. Finally, show that if two Hermitian matrices have the same inertia matrix, then they must be congruent. **Hint:** First recall that there is a unitary matrix, U such that

$$U^*AU = \begin{pmatrix} D_{n_+} & & \\ & D_{n_-} & \\ & & O_{n_0} \end{pmatrix}$$

where the D_{n_+} is a diagonal matrix having the positive eigenvalues of A, D_{n_-} being defined similarly. Now let $|D_{n_-}|$ denote the diagonal matrix which replaces each entry of D_{n_-} with its absolute value. Consider the two diagonal matrices

$$D = D^* = \begin{pmatrix} D_{n_+}^{-1/2} & & \\ & |D_{n_-}|^{-1/2} & \\ & & I_{n_0} \end{pmatrix}$$

Now consider D^*U^*AUD.

44. Show that if A, B are two congruent Hermitian matrices, then they have the same inertia matrix. **Hint:** Let $A = SBS^*$ where S is invertible. Show that A, B have the same rank and this implies that they are each unitarily similar to a diagonal matrix which has the same number of zero entries on the main diagonal. Therefore, letting V_A be the span of the eigenvectors associated with positive eigenvalues of A and V_B being defined similarly, it suffices to show that these have the same dimensions. Show that $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \in V_A$. Next consider S^*V_A. For $\mathbf{x} \in V_A$, explain why

$$\begin{aligned} (BS^*\mathbf{x}, S^*\mathbf{x}) &= \left(S^{-1}A(S^*)^{-1}S^*\mathbf{x}, S^*\mathbf{x}\right) \\ &= \left(S^{-1}A\mathbf{x}, S^*\mathbf{x}\right) = \left(A\mathbf{x}, \left(S^{-1}\right)^* S^*\mathbf{x}\right) = (A\mathbf{x}, \mathbf{x}) > 0 \end{aligned}$$

Next explain why this shows that $S^* V_A$ is a subspace of V_B and so the dimension of V_B is at least as large as the dimension of V_A. Hence there are at least as many positive eigenvalues for B as there are for A. Switching A, B you can turn the inequality around. Thus the two have the same inertia matrix.

45. Let A be an $m \times n$ matrix. Then if you unraveled it, you could consider it as a vector in $\mathbb{C}^{nm}$. The Frobenius inner product on the vector space of $m \times n$ matrices is defined as
$$(A, B) \equiv \text{trace}(AB^*)$$
Show that this really does satisfy the axioms of an inner product space and that it also amounts to nothing more than considering $m \times n$ matrices as vectors in $\mathbb{C}^{nm}$.

46. ↑Consider the $n \times n$ unitary matrices. Show that whenever U is such a matrix, it follows that
$$|U|_{\mathbb{C}^{nn}} = \sqrt{n}$$
Next explain why if $\{U_k\}$ is any sequence of unitary matrices, there exists a subsequence $\{U_{k_m}\}_{m=1}^{\infty}$ such that $\lim_{m\to\infty} U_{k_m} = U$ where U is unitary. Here the limit takes place in the sense that the entries of U_{k_m} converge to the corresponding entries of U.

47. ↑Let A, B be two $n \times n$ matrices. Denote by $\sigma(A)$ the set of eigenvalues of A. Define
$$\text{dist}(\sigma(A), \sigma(B)) = \max_{\lambda \in \sigma(A)} \min\{|\lambda - \mu| : \mu \in \sigma(B)\}$$
Explain why $\text{dist}(\sigma(A), \sigma(B))$ is small if and only if every eigenvalue of A is close to some eigenvalue of B. Now prove the following theorem using the above problem and Schur's theorem. This theorem says roughly that if A is close to B then the eigenvalues of A are close to those of B in the sense that every eigenvalue of A is close to an eigenvalue of B.

Theorem 7.10.2 *Suppose* $\lim_{k\to\infty} A_k = A$. *Then*
$$\lim_{k\to\infty} \text{dist}(\sigma(A_k), \sigma(A)) = 0$$

48. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2×2 matrix which is not a multiple of the identity. Show that A is similar to a 2×2 matrix which has at least one diagonal entry equal to 0. **Hint:** First note that there exists a vector $\mathbf{a}$ such that $A\mathbf{a}$ is not a multiple of $\mathbf{a}$. Then consider
$$B = \begin{pmatrix} \mathbf{a} & A\mathbf{a} \end{pmatrix}^{-1} A \begin{pmatrix} \mathbf{a} & A\mathbf{a} \end{pmatrix}$$
Show B has a zero on the main diagonal.

49. ↑ Let A be a complex $n \times n$ matrix which has trace equal to 0. Show that A is similar to a matrix which has all zeros on the main diagonal. **Hint:** Use Problem 30 on Page 158 to argue that you can say that a given matrix is similar to one which has the diagonal entries permuted in any order desired. Then use the above problem and block multiplication to show that if the A has k nonzero entries, then it is similar to a matrix which has $k-1$ nonzero entries. Finally, when A is similar to one which has at most one nonzero entry, this one must also be zero because of the condition on the trace.

50. ↑An $n \times n$ matrix X is a comutator if there are $n \times n$ matrices A, B such that $X = AB - BA$. Show that the trace of any comutator is 0. Next show that if a complex matrix X has trace equal to 0, then it is in fact a comutator. **Hint:** Use the above problem to show that it suffices to consider X having all zero entries on the main diagonal. Then define

$$A = \begin{pmatrix} 1 & & & 0 \\ & 2 & & \\ & & \ddots & \\ 0 & & & n \end{pmatrix},\ B_{ij} = \begin{cases} \frac{X_{ij}}{i-j} \text{ if } i \neq j \\ 0 \text{ if } i = j \end{cases}$$

Vector Spaces And Fields

8.1 Vector Space Axioms

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

Definition 8.1.1 *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.)

$$\alpha(\mathbf{v}+\mathbf{w}) = \alpha\mathbf{v}+\alpha\mathbf{w}, \tag{8.1}$$

$$(\alpha+\beta)\mathbf{v} = \alpha\mathbf{v}+\beta\mathbf{v}, \tag{8.2}$$

$$\alpha(\beta\mathbf{v}) = \alpha\beta(\mathbf{v}), \tag{8.3}$$

$$1\mathbf{v} = \mathbf{v}. \tag{8.4}$$

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 8.1.2 *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 8.1.3 *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.*

8.2 Subspaces And Bases

8.2.1 Basic Definitions

Definition 8.2.1 *If $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\} \subseteq V$, a vector space, then*

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

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 $W \subseteq V$ 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 8.2.2 *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 8.2.3 *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

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

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

8.2.2 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 8.2.4 *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 $\text{span}\{\mathbf{y}_1, \cdots, \mathbf{y}_s\} \equiv V$, it follows there exist scalars $c_1, \cdots, c_s$ such that

$$\mathbf{x}_1 = \sum_{i=1}^{s} c_i \mathbf{y}_i. \tag{8.5}$$

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 8.5 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

$$\mathbf{x}_{l+1} = \sum_{i=1}^{l} c_i \mathbf{x}_i + \sum_{j=1}^{p} d_j \mathbf{z}_j. \tag{8.6}$$

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, (8.6) 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

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

But then $\mathbf{x}_r \in \operatorname{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_s\}$ 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 8.2.5 *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 8.2.4, $m \leq n$ and $n \leq m$. ■

Definition 8.2.6 *A vector space V is of dimension n if it has a basis consisting of n vectors. This is well defined thanks to Corollary 8.2.5. It is always assumed here that $n < \infty$ and in this case, such a vector space is said to be finite dimensional.*

Example 8.2.7 *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 8.2.8 *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 8.2.9 *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 8.2.10 *Suppose $\mathbf{v} \notin \operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k)$ 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 8.2.11 *If $V = \operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_n)$ 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

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

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

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

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

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

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

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

Then by Lemma 8.2.10, $\{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ 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 8.2.4. Therefore, this list is a basis. ■

8.2.3 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 8.2.12 *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 $\operatorname{span}\{\mathbf{v}_1\} = V$, stop. $\{\mathbf{v}_1\}$ is a basis for V. Otherwise, there exists $\mathbf{v}_2 \in V$ which is not in $\operatorname{span}\{\mathbf{v}_1\}$. By Lemma 8.2.10 $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set of vectors. If $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} = V$ stop, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for V. If $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} \neq V$, then there exists $\mathbf{v}_3 \notin \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ 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 8.2.4. ■

8.3 Lots Of Fields

8.3.1 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 8.3.1 *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.*

Example 8.3.2 *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.*

Lemma 8.3.3 *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 $r(\lambda) = 0$.

Proof: Suppose that $f(\lambda) - q(\lambda) g(\lambda)$ is never equal to 0 for any $q(\lambda)$. If it is, then the conclusion follows.

Denote by S the set of polynomials $f(\lambda) - g(\lambda) l(\lambda)$. Out of all these polynomials, there exists one which has smallest degree $r(\lambda)$. Let this take place when $l(\lambda) = q_1(\lambda)$. Thus

$$r(\lambda) = f(\lambda) - g(\lambda) q_1(\lambda).$$

By assumption, $r(\lambda) \neq 0$. It is required to show the degree of $r(\lambda)$ is smaller than the degree of $g(\lambda)$. If this doesn't happen, then the degree of $r(\lambda) \geq$ the degree of $g(\lambda)$. Let

$$\begin{aligned} r(\lambda) &= b_m \lambda^m + \cdots + b_1 \lambda + b_0 \\ g(\lambda) &= a_n \lambda^n + \cdots + a_1 \lambda + a_0 \end{aligned}$$

where $m \geq n$ and b_m and a_n are nonzero. Then let $r_1(\lambda)$ be given by

$$r_1(\lambda) = r(\lambda) - \frac{\lambda^{m-n} b_m}{a_n} g(\lambda)$$

Thus

$$r_1(\lambda) = (b_m \lambda^m + \cdots + b_1 \lambda + b_0) - \frac{\lambda^{m-n} b_m}{a_n} (a_n \lambda^n + \cdots + a_1 \lambda + a_0)$$

which has degree at most $m - 1$. But

$$\begin{aligned} r_1(\lambda) &= \overbrace{f(\lambda) - g(\lambda) q_1(\lambda)}^{r(\lambda)} - \frac{\lambda^{m-n} b_m}{a_n} g(\lambda) \\ &= f(\lambda) - g(\lambda) \left(q_1(\lambda) + \frac{\lambda^{m-n} b_m}{a_n} \right), \end{aligned}$$

and this is one of the polynomials in S, contradicting the definition of $r(\lambda)$ which required it has the smallest degree. ■

In fact, the polynomials $r(\lambda), q(\lambda)$ are unique. Suppose $(r(\lambda), q(\lambda))$, and $(r_1(\lambda), q_1(\lambda))$ are two pairs which work. Then

$$(q_1(\lambda) - q(\lambda)) g(\lambda) = r(\lambda) - r_1(\lambda)$$

The degree of the polynomial on the right, would need to be less than the degree of the polynomial on the left if it is not zero. Hence $r_1(\lambda) = r(\lambda)$. Now one can argue, by comparing coefficients, that $q_1(\lambda) = q(\lambda)$.

Now with this lemma, here is another one which is very fundamental. First here is a definition. A polynomial is **monic** means it is of the form

$$\lambda^n + c_{n-1} \lambda^{n-1} + \cdots + c_1 \lambda + c_0.$$

That is, the leading coefficient is 1. 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 8.3.4 *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.*

Proposition 8.3.5 *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 8.3.6 *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 8.3.3

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

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 8.3.7 *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 8.3.8 *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. Then $\psi(\lambda)$ equals some $\phi_i(\lambda)$.

Proof : Suppose $\psi(\lambda) \neq \phi_i(\lambda)$ for all i. Then by Lemma 8.3.7, there exist polynomials $m_i(\lambda), n_i(\lambda)$ such that

$$1 = \psi(\lambda) m_i(\lambda) + \phi_i(\lambda) n_i(\lambda).$$

Hence

$$(\phi_i(\lambda) n_i(\lambda))^{k_i} = (1 - \psi(\lambda) m_i(\lambda))^{k_i}$$

Then, letting $\widetilde{g}(\lambda) = \prod_{i=1}^{p} n_i(\lambda)^{k_i}$, and applying the binomial theorem, there exists a polynomial $h(\lambda)$ such that

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

Thus, using the fact that $\psi(\lambda)$ divides $\prod_{i=1}^{p} \phi_i(\lambda)^{k_i}$, for a suitable polynomial $g(\lambda)$,

$$g(\lambda) \psi(\lambda) = 1 + \psi(\lambda) h(\lambda)$$

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

which is impossible if $\psi(\lambda)$ is non constant, as assumed. ■

Now here is a simple lemma about canceling monic polynomials.

Lemma 8.3.9 *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

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

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

Theorem 8.3.10 *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 8.3.8, each $\psi_i(\lambda)$ equals one of the $\phi_j(\lambda)$. By the above cancellation lemma, Lemma 8.3.9, 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 $a = b$. $\blacksquare$

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

Corollary 8.3.11 *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 8.3.10, 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. ■

8.3.2 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 8.3.12 *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 8.3.13 *$[x]$ 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 8.3.14 *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 8.3.15 *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 8.3.16 *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 8.3.17 *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. For the similarity relation defined in Definition 8.3.15, 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 8.3.18 *In the situation of Definition 8.3.17, $p(x)$ and $q(x)$ are relatively prime for any $q(x) \in \mathbb{F}[x]$ which is not a multiple of $p(x)$. Also the definitions of addition and multiplication are well defined. In addition, if $a, b \in \mathbb{F}$ and $[a] = [b]$, then $a = b$.*

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 $p(x)$. 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. ■

Note that from this proposition and math induction, if each $a_i \in \mathbb{F}$,

$$\begin{aligned} & \left[a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0\right] \\ = \;& [a_n]\,[x]^n + [a_{n-1}]\,[x]^{n-1} + \cdots [a_1]\,[x] + [a_0] \end{aligned} \tag{8.7}$$

With the above preparation, here is a definition of a field in which the irreducible polynomial $p(x)$ has a root.

Definition 8.3.19 *Let $p(x) \in \mathbb{F}[x]$ be irreducible and let $a(x) \sim b(x)$ when $a(x) - b(x)$ is a multiple of $p(x)$. Let $\mathbb{G}$ denote the set of equivalence classes as described above with the operations also described in Definition 8.3.17.*

Also here is another useful definition and a simple proposition which comes from it.

Definition 8.3.20 *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 8.3.21 *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. ■

Theorem 8.3.22 *The set of all equivalence classes $\mathbb{G} \equiv \mathbb{F}/(p(x))$ described above with the multiplicative identity given by $[1]$ and the additive identity given by $[0]$ along with the operations of Definition 8.3.17, is a field and $p([x]) = [0]$. (Thus p has a root in this new field.) In addition to this, $[\mathbb{G} : \mathbb{F}] = n$, the degree of $p(x)$.*

Proof: Everything is obvious except for the existence of the multiplicative inverse and the assertion that $p([x]) = 0$. Suppose then that $[a(x)] \neq [0]$. That is, $a(x)$ is not a multiple of $p(x)$. Why does $[a(x)]^{-1}$ exist? By Theorem 8.3.6, $a(x), p(x)$ are relatively prime and so there exist polynomials $\psi(x), \phi(x)$ such that

$$1 = \psi(x) p(x) + a(x) \phi(x)$$

and so

$$1 - a(x)\phi(x) = \psi(x)p(x)$$

which, by definition implies

$$[1 - a(x)\phi(x)] = [1] - [a(x)\phi(x)] = [1] - [a(x)][\phi(x)] = [0]$$

and so $[\phi(x)] = [a(x)]^{-1}$. This shows $\mathbb{G}$ is a field.

Now if $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, $p([x]) = 0$ by 8.7 and the definition which says $[p(x)] = [0]$.

Consider the claim about the dimension. It was just shown that $[1], [x], [x^2], \cdots, [x^n]$ is linearly dependent. Also $[1], [x], [x^2], \cdots, [x^{n-1}]$ is independent because if not, there would exist a polynomial $q(x)$ of degree $n-1$ which is a multiple of $p(x)$ which is impossible. Now for $[q(x)] \in \mathbb{G}$, you can write

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

where the degree of $r(x)$ is less than n or else it equals 0. Either way, $[q(x)] = [r(x)]$ which is a linear combination of $[1], [x], [x^2], \cdots, [x^{n-1}]$. Thus $[\mathbb{G} : \mathbb{F}] = n$ as claimed. ■

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

Usually, people simply 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.

Theorem 8.3.23 *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 $\mathbb{G}$ such that there exist $\{z_1, \cdots, z_n\}$ 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 Theorem 8.3.22, 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 Theorem 8.3.22, there is a larger field $\mathbb{G}_1$ such that $p(x)$ has a root a_1 in $\mathbb{G}_1$ and $[\mathbb{G} : \mathbb{F}] \leq n$. Then

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

Continue this way until the polynomial equals the product of linear factors. Then by Proposition 8.3.21 applied multiple times, $[\mathbb{G} : \mathbb{F}] \leq n!$. ■

Example 8.3.24 *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 8.3.25 *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 8.3.26 *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 $(ai+b)^{-1} = \frac{b-ai}{a^2+b^2}$.

8.3.3 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 8.3.27 *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* $p(a) = 0$*.*

Theorem 8.3.28 *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 the 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 and irreducible such that $q(a) = 0$.

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

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 $q(x)$ is irreducible, $l(x) = 1$ showing that $p(x) = q(x)$. ■

Definition 8.3.29 *For* a *an algebraic number, let* $\deg(a)$ *denote the degree of the minimal polynomial of* a*.*

Also, here is another definition.

Definition 8.3.30 *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 8.3.31 *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 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. Thus $\mathbb{F}[a]$ is spanned by

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

Since $p(x)$ has smallest degree of all polynomial 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 $\left\{a_1^{k_1}, a_2^{k_2}, \cdots, a_n^{k_n}\right\}$ 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 8.2.4 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 8.3.32 *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 296 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 8.3.33 *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 8.3.31 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 $\left(x - \beta_j\right)$ for some β_j for some $j > 1$ because these are the only factors of $g(x)$. Therefore, this would require

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

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.

8.3.4 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 8.3.34 *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.

8.4 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 $M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_3 = u_1 = 0 \right\}$. Is M a subspace? Explain.

3. Let $M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_3 \geq u_1 \right\}$. Is M a subspace? Explain.

4. Let $\mathbf{w} \in \mathbb{R}^4$ and let $M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : \mathbf{w} \cdot \mathbf{u} = 0 \right\}$. Is M a subspace? Explain.

5. Let $M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_i \geq 0 \text{ for each } i = 1, 2, 3, 4 \right\}$. Is M a subspace? Explain.

6. Let $\mathbf{w}, \mathbf{w}_1$ be given vectors in $\mathbb{R}^4$ and define
$$M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : \mathbf{w} \cdot \mathbf{u} = 0 \text{ and } \mathbf{w}_1 \cdot \mathbf{u} = 0 \right\}.$$

Is M a subspace? Explain.

7. Let $M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : |u_1| \leq 4 \right\}$. Is M a subspace? Explain.

8. Let $M = \left\{ \mathbf{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : \sin(u_1) = 1 \right\}$. 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 $\text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_k)$.

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 265.

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.

Linear Transformations

9.1 Matrix Multiplication As A Linear Transformation

Definition 9.1.1 *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

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

9.2 $\mathcal{L}(V,W)$ As A Vector Space

Definition 9.2.1 *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.

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

Lemma 9.2.2 *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 9.2.3 *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 9.2.2, 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

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

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

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

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. ■

9.3 The Matrix Of A Linear Transformation

Definition 9.3.1 *In Theorem 9.2.3, 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.

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 9.3.2 *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.

$$\begin{array}{ccccc} & & L & & \\ \beta = \{v_1, \cdots, v_n\} & V & \rightarrow & W & \{w_1, \cdots, w_m\} = \gamma \\ & q_\beta \uparrow & \circ & \uparrow q_\gamma & \\ & \mathbb{F}^n & \rightarrow & \mathbb{F}^m & \\ & & [L]_{\gamma\beta} & & \end{array} \tag{9.1}$$

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

$$\begin{pmatrix} Lv_1 & \cdots & Lv_n \end{pmatrix} = \begin{pmatrix} w_1 & \cdots & w_m \end{pmatrix} [L]_{\gamma\beta} \tag{9.2}$$

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 9.3.3 *Let*

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

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

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 9.2,

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

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 9.3.4 *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 9.3.5 *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 $\mathbb{F}^n$,

$$S^{-1} A_2 S = A_1. \tag{9.3}$$

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 9.3.6 *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 9.3.7 *$[x]$ 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 9.3.8 *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 9.3.9 *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, if $A \sim B$, let L be as just described. Thus $L = q_\beta A q_\beta^{-1} = q_\beta SBS^{-1}q_\beta^{-1}$. Let $q_\gamma \equiv q_\beta S$ and it follows that B is the matrix of L with respect to $\{q_\beta S\mathbf{e}_1, \cdots, q_\beta S\mathbf{e}_n\}$. ■

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 9.3.10 *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.

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

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 9.3.11 *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 9.3.12 *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 9.3.11 consists of the $n \times n$ matrix whose columns are the eigenvectors of A and $D = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$.*

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 $S^{-1}AS = D \equiv \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$. 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 9.3.13 *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 9.3.9 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 9.3.14 *Let $A \in \mathcal{L}(X, Y)$ where X and Y are finite dimensional vector spaces. Define* $\operatorname{rank}(A)$ *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 9.3.15 *Let $A \in \mathcal{L}(X, Y)$. Then* $\operatorname{rank}(A) = \operatorname{rank}(M)$ *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.

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

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 9.3.16 *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.

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

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

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

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

9.3.1 Some Geometrically Defined Linear Transformations

If T is any linear transformation which maps $\mathbb{F}^n$ to $\mathbb{F}^m$, there is always an $m \times n$ matrix $A \equiv [T]$ with the property that

$$A\mathbf{x} = T\mathbf{x} \tag{9.4}$$

for all $\mathbf{x} \in \mathbb{F}^n$. You simply take the matrix of the linear transformation with respect to the standard basis. What is the form of A? Suppose $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation and you want to find the matrix defined by this linear transformation as described in 9.4. Then if $\mathbf{x} \in \mathbb{F}^n$ it follows

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$

where $\mathbf{e}_i$ is the vector which has zeros in every slot but the i^{th} and a 1 in this slot. Then since T is linear,

$$T\mathbf{x} = \sum_{i=1}^{n} x_i T(\mathbf{e}_i)$$

$$= \begin{pmatrix} | & & | \\ T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \equiv A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

and so you see that the matrix desired is obtained from letting the i^{th} column equal $T(\mathbf{e}_i)$. This proves the following theorem.

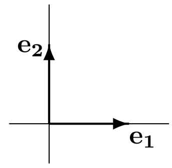
Theorem 9.3.18 *Let T be a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$. Then the matrix A satisfying 9.4 is given by*

$$\begin{pmatrix} | & & | \\ T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \\ | & & | \end{pmatrix}$$

where $T\mathbf{e}_i$ is the i^{th} column of A.

Example 9.3.19 *Determine the matrix for the transformation mapping $\mathbb{R}^2$ to $\mathbb{R}^2$ which consists of rotating every vector counter clockwise through an angle of θ.*

Let $\mathbf{e}_1 \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{e}_2 \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. These identify the geometric vectors which point along the positive x axis and positive y axis as shown.



From Theorem 9.3.18, you only need to find $T\mathbf{e}_1$ and $T\mathbf{e}_2$, the first being the first column of the desired matrix A and the second being the second column. From drawing a picture and doing a little geometry, you see that

$$T\mathbf{e}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, T\mathbf{e}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

Therefore, from Theorem 9.3.18,

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example 9.3.20 *Find the matrix of the linear transformation which is obtained by first rotating all vectors through an angle of ϕ and then through an angle θ. Thus you want the linear transformation which rotates all angles through an angle of $\theta + \phi$.*

Let $T_{\theta+\phi}$ denote the linear transformation which rotates every vector through an angle of $\theta + \phi$. Then to get $T_{\theta+\phi}$, you could first do T_ϕ and then do T_θ where T_ϕ is the linear transformation which rotates through an angle of ϕ and T_θ is the linear transformation which rotates through an angle of θ. Denoting the corresponding matrices by $A_{\theta+\phi}$, A_ϕ, and A_θ, you must have for every $\mathbf{x}$

$$A_{\theta+\phi}\mathbf{x} = T_{\theta+\phi}\mathbf{x} = T_\theta T_\phi \mathbf{x} = A_\theta A_\phi \mathbf{x}.$$

Consequently, you must have

$$\begin{aligned} A_{\theta+\phi} &= \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix} = A_\theta A_\phi \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}. \end{aligned}$$

Therefore,

$$\begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & \cos\theta\cos\phi - \sin\theta\sin\phi \end{pmatrix}.$$

Don't these look familiar? They are the usual trig. identities for the sum of two angles derived here using linear algebra concepts.

Example 9.3.21 *Find the matrix of the linear transformation which rotates vectors in* $\mathbb{R}^3$ *counterclockwise about the positive* z *axis.*

Let T be the name of this linear transformation. In this case, $T\mathbf{e}_3 = \mathbf{e}_3, T\mathbf{e}_1 = (\cos\theta, \sin\theta, 0)^T$, and $T\mathbf{e}_2 = (-\sin\theta, \cos\theta, 0)^T$. Therefore, the matrix of this transformation is just

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.5}$$

In Physics it is important to consider the work done by a force field on an object. This involves the concept of projection onto a vector. Suppose you want to find the projection of a vector, $\mathbf{v}$ onto the given vector, $\mathbf{u}$, denoted by $\text{proj}_{\mathbf{u}}(\mathbf{v})$ This is done using the dot product as follows.

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}$$

Because of properties of the dot product, the map $\mathbf{v} \to \text{proj}_{\mathbf{u}}(\mathbf{v})$ is linear,

$$\begin{aligned} \text{proj}_{\mathbf{u}}(\alpha\mathbf{v}+\beta\mathbf{w}) &= \left(\frac{\alpha\mathbf{v}+\beta\mathbf{w}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u} = \alpha\left(\frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u} + \beta\left(\frac{\mathbf{w}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u} \\ &= \alpha\,\text{proj}_{\mathbf{u}}(\mathbf{v}) + \beta\,\text{proj}_{\mathbf{u}}(\mathbf{w}). \end{aligned}$$

Example 9.3.22 *Let the projection map be defined above and let $\mathbf{u} = (1,2,3)^T$. Find the matrix of this linear transformation with respect to the usual basis.*

You can find this matrix in the same way as in earlier examples. $\text{proj}_{\mathbf{u}}(\mathbf{e}_i)$ gives the i^{th} column of the desired matrix. Therefore, it is only necessary to find

$$\text{proj}_{\mathbf{u}}(\mathbf{e}_i) \equiv \left(\frac{\mathbf{e}_i\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}$$

For the given vector in the example, this implies the columns of the desired matrix are

$$\frac{1}{14}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \frac{2}{14}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \frac{3}{14}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Hence the matrix is

$$\frac{1}{14}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}.$$

Example 9.3.23 *Find the matrix of the linear transformation which reflects all vectors in $\mathbb{R}^3$ through the xz plane.*

As illustrated above, you just need to find $T\mathbf{e}_i$ where T is the name of the transformation. But $T\mathbf{e}_1 = \mathbf{e}_1, T\mathbf{e}_3 = \mathbf{e}_3$, and $T\mathbf{e}_2 = -\mathbf{e}_2$ so the matrix is

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

Example 9.3.24 *Find the matrix of the linear transformation which first rotates counter clockwise about the positive z axis and then reflects through the xz plane.*

This linear transformation is just the composition of two linear transformations having matrices

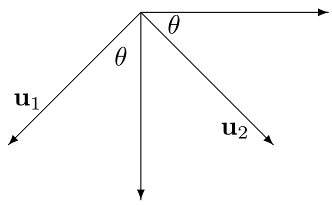
$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

respectively. Thus the matrix desired is

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

9.3.2 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 9.3.10, 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

$$= \begin{pmatrix} a^2 + \left(1-a^2\right)\cos\theta & ab(1-\cos\theta) - c\sin\theta & ac(1-\cos\theta) + b\sin\theta \\ ab(1-\cos\theta) + c\sin\theta & b^2 + \left(1-b^2\right)\cos\theta & bc(1-\cos\theta) - a\sin\theta \\ ac(1-\cos\theta) - b\sin\theta & bc(1-\cos\theta) + a\sin\theta & c^2 + \left(1-c^2\right)\cos\theta \end{pmatrix}. \tag{9.6}$$

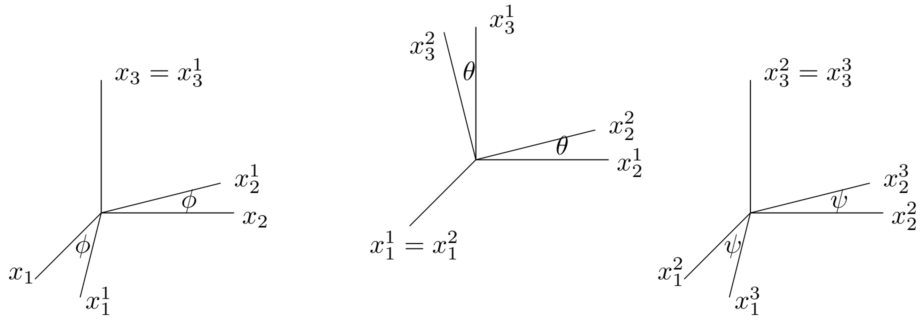
With this, it is clear how to rotate clockwise about the unit vector, (a, b, c). Just rotate counter clockwise through an angle of $-\theta$. Thus the matrix for this clockwise rotation is just

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

In deriving 9.6 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 9.6 reduces to 9.5.

9.3.3 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 9.6, $a = \cos\phi, b = \sin\phi$, 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

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

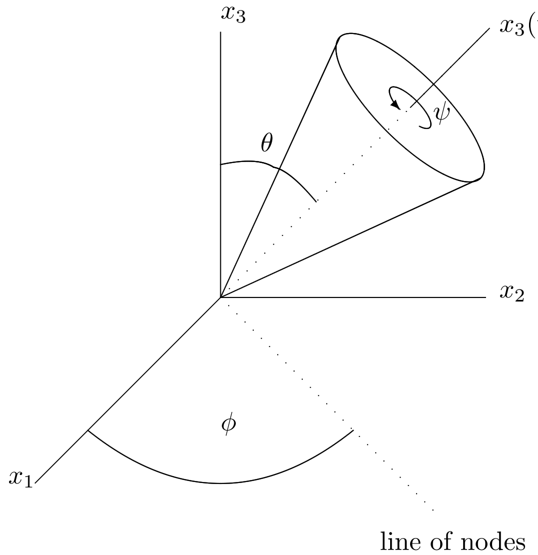
and it is desired to rotate counter clockwise through an angle of ψ about this vector. Thus, in this case,

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

and you could substitute in to the formula of Theorem 9.6 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 9.3.17, 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.

9.4 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 9.4.1 *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 9.4.2 *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 9.3.16. 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 9.4.3 *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 9.2.3 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,

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

Since $\tilde{q}(A) = 0$, 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 9.4.4 *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 $\mathbb{F}$.

9.5 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. $\uparrow$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 $[T]_\gamma =$
$$\begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} [T]_\beta \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}^{-1} \tag{9.7}$$

10. $\uparrow$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 $\gamma = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, 9.7 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 [23]. **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 9.2.2 verify that L is linear.

Canonical Forms

10.1 A Theorem Of Sylvester, Direct Sums

The notation is defined as follows.

Definition 10.1.1 *Let* $L \in \mathcal{L}(V, W)$. *Then* $\ker(L) \equiv \{v \in V : Lv = 0\}$.

Lemma 10.1.2 *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

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

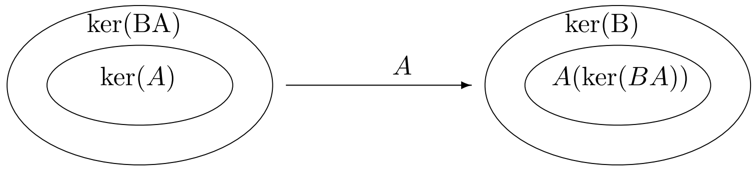
Suppose now that $A \in \mathcal{L}(V, W)$ and $B \in \mathcal{L}(W, U)$ where V, W, U are all finite dimensional vector spaces. Then it is interesting to consider $\ker(BA)$. The following theorem of Sylvester is a very useful and important result.

Theorem 10.1.3 *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 $Az = \sum_{i=1}^{m} a_i A y_i$ 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 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 10.1.4 *Let $\{V_i\}_{i=1}^r$ be subspaces of V. Then*

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

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

$$\sum_{i=1}^{r} v_i = 0, v_i \in V_i, \tag{10.1}$$

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 10.1.5 *Let* $L \in \mathcal{L}(V, W)$ *where* V, W *are* n *dimensional vector spaces. Then if* L *is one to one, it follows that* 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 8.2.10, $\{Lv_1, \cdots, Lv_n, w\}$ would be independent and this contradicts the exchange theorem, Theorem 8.2.4 because it would be a linearly independent set having more vectors than the spanning set $\{v_1, \cdots, v_n\}$. ■

Lemma 10.1.6 *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 10.1.7 *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: First suppose $p = 2$. Then denote the two operators by A, B respectively. By Sylvester,

$$\dim(BA) \leq \dim(A) + \dim(B) \tag{10.2}$$

Does equality hold? By Lemma 10.1.3 it suffices to check whether $A(\ker(BA)) = \ker(B)$. By Lemma 10.1.5, and since the operators commute,

$$A : \ker(B) \to \ker(B), \text{ one to one and onto.}$$

If $y \in \ker(B)$, is $y = Az$ where $z \in \ker(BA)$? Let $z = A^{-1}y$. Then $z \in \ker(B)$ and

$$BAz = BAA^{-1}y = By = 0.$$

Thus equality holds in the above Sylvester inequality 10.2. Now if $a \in \ker(A)$ and $b \in \ker(B)$, and $a + b = 0$, then

$$0 = A(a + b) = Aa + Ab = Ab$$

Since $Ab = 0$, and A is one to one on $\ker(B)$, it follows that $b = 0$. Similarly $a = 0$ and so, since these operators commute,

$$\ker(A) + \ker(B) = \ker(A) \oplus \ker(B) \subseteq \ker(BA)$$

Then it follows from Sylvester's inequality again that

$$\begin{aligned} \dim(\ker(BA)) &\leq \dim(\ker(A)) + \dim(\ker(B)) \\ &= \dim(\ker(A) \oplus \ker(B)) \leq \dim(\ker(BA)) \end{aligned}$$

and so $\ker(BA) = \ker(A) \oplus \ker(B)$. Thus the lemma is true if $p = 2$. Suppose it is true for $p - 1$. Then from what was just shown and induction,

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

10.2 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 9.4.3 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 10.2.1 *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 10.2.2 *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 10.2.3 *In the context of Definition 10.2.2,*

$$V = V_1 \oplus \cdots \oplus V_q \tag{10.3}$$

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 10.1.7 and the fact that $\ker\left(\prod_{k=1}^{q}\phi_k(\lambda)^{r_k}\right) = V$, 10.3 is obtained. The claim about the bases follows from Lemma 10.1.6. ■

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 10.2.4 *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 8.3.11 $\eta(\lambda) = \phi_k(\lambda)^{r_k}$ 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 10.2.3, 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 10.2.3 relate to matrices?

Theorem 10.2.5 *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 $q_k : F^{|\beta_k|} \to V_k$ be the usual map such that the following diagram commutes.

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

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,

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

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 10.2.6 *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.

10.3 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 10.3.1 *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.

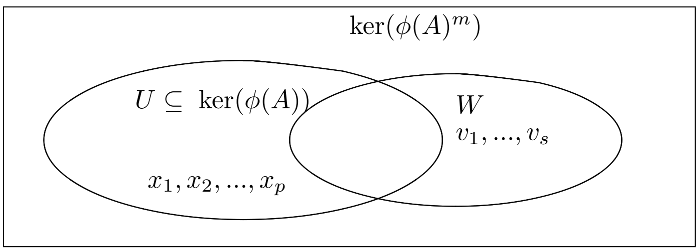
Lemma 10.3.2 *Let $\beta_x = \left\{x, Ax, A^2x, \cdots, A^{m-1}x\right\}, x \neq 0$ where m is the smallest such that $A^m x \in \operatorname{span}\left(x, \cdots, A^{m-1}x\right)$. Then β_x is linearly independent.*

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. ■

For more on the next lemma and the following theorem, see [14]. I am following the presentation in Friedberg Insel and Spence [9]. See also Herstein [13]. To help organize the ideas in the lemma, here is a diagram.



Lemma 10.3.3 *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 $|\beta_x| = d$ and so

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

also $\text{span}(\beta_x)$ 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 $\text{span}(\beta_x)$ is obvious because $A^d x \in \text{span}(\beta_x)$. 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 $W, \text{span}\left(x, Ax, \cdots, A^{d-1}x\right)$ are A invariant. Therefore,

$$\begin{aligned} \text{span}\left(z, Az, \cdots, A^{d-1}z\right) &\subseteq W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right) \\ &\subseteq \text{span}\left(x, Ax, \cdots, A^{d-1}x\right) \end{aligned} \tag{10.4}$$

Suppose $z \neq 0$. Then from the Lemma 10.3.2 above, $\left\{z, Az, \cdots, A^{d-1}z\right\}$ must be linearly independent. Therefore,

$$d = \dim\left(\text{span}\left(z, Az, \cdots, A^{d-1}z\right)\right) \leq \dim\left(W \cap \text{span}\left(x, Ax, \cdots, A^{d-1}x\right)\right)$$

$$\leq \dim\left(\operatorname{span}\left(x, Ax, \cdots, A^{d-1}x\right)\right) = d$$

Thus

$$W \cap \operatorname{span}\left(x, Ax, \cdots, A^{d-1}x\right) = \operatorname{span}\left(x, Ax, \cdots, A^{d-1}x\right)$$

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 10.3.2, 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 $\operatorname{span}\left(v_1, \cdots, v_s, \beta_{x_1}, \cdots, \beta_{x_n}\right) = U + W$. The process stops because $\ker\left(\phi(A)^m\right)$ is finite dimensional.

Finally, letting $x \in \ker\left(\phi(A)^m\right)$, 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 8.3.11 $\eta(\lambda) = \phi(\lambda)^k$ 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 [14] which pertains to the existence of complementary subspaces. For an outline of this generalization, see Problem 9 on 404.

Theorem 10.3.4 *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. Then there exists a basis for V which is of the form $\beta = \left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$.*

Proof: First suppose $m = 1$. Then in Lemma 10.3.3 you can let $W = \{0\}$ and $U = \ker(\phi(A))$. Then by this lemma, there exist $v_1, v_2, \cdots, v_s$ such that $\left\{\beta_{v_1}, \cdots, \beta_{v_s}\right\}$ is a basis for $\ker(\phi(A))$. Suppose then that the theorem is true for $m - 1$, $m \geq 2$.

Now let the minimal polynomial for A on V be $\phi(A)^m$ where $\phi(\lambda)$ is monic and irreducible. Then $\phi(A)(V)$ is an invariant subspace of V. What is the minimal polynomial of A on $\phi(A)(V)$? Clearly $\phi(A)^{m-1}$ will send everything in $\phi(A)(V)$ to 0. If $\eta(\lambda)$ is the minimal polynomial of A on $\phi(A)(V)$, then

$$\phi(\lambda)^{m-1} = l(\lambda)\eta(\lambda) + r(\lambda)$$

and $r(\lambda)$ must equal 0 since otherwise $r(A) = 0$ and $\eta(\lambda)$ was not minimal. By Corollary 8.3.11, $\eta(\lambda) = \phi(\lambda)^k$ for some $k \leq m - 1$. However, it cannot happen that $k < m - 1$ because if so, $\phi(\lambda)^m$ would fail to be the minimal polynomial for A on V. By induction, $\phi(A)(V)$ has a basis $\left\{\beta_{x_1}, \cdots, \beta_{x_p}\right\}$.

Let $y_j \in V$ be such that $\phi(A) y_j = x_j$. Consider $\left\{\beta_{y_1}, \cdots, \beta_{y_p}\right\}$. Are these vectors independent? Suppose

$$0 = \sum_{i=1}^{p} \sum_{j=1}^{|\beta_{y_i}|} a_{ij} A^{j-1} y_i \equiv \sum_{i=1}^{p} f_i(A) y_i \tag{10.5}$$

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

$$0 = \sum_{i=1}^{p} \sum_{j=1}^{|\beta_{y_i}|} a_{ij} A^{j-1} x_i \equiv \sum_{i=1}^{p} f_i(A) x_i$$

Now $f_i(A) x_i = 0$ for each i since $f_i(A) x_i \in \operatorname{span}\left(\beta_{x_i}\right)$. Let $\eta_i(\lambda)$ be the monic polynomial of smallest degree such that $\eta_i(A) x_i = 0$. It follows from the usual division algorithm that $\eta_i(\lambda)$ divides $f_i(\lambda)$. Also, $\phi(A)^{m-1} x_i = 0$ and so $\eta_i(\lambda)$ must divide $\phi(\lambda)^{m-1}$. From

Corollary 8.3.11, it follows that, since $\phi(\lambda)$ is irreducible, $\eta_i(\lambda) = \phi(\lambda)^k$ for some $k \leq m-1$. Thus $\phi(\lambda)$ divides $\eta_i(\lambda)$ which divides $f_i(\lambda)$. Hence $f_i(\lambda) = \phi(\lambda) g_i(\lambda)$. Now

$$0 = \sum_{i=1}^{p} f_i(A) y_i = \sum_{i=1}^{p} g_i(A) \phi(A) y_i = \sum_{i=1}^{p} g_i(A) x_i.$$

By the same reasoning just given, since $g_i(A) x_i \in \operatorname{span}(\beta_{x_i})$, it follows that each $g_i(A) x_i = 0$. Therefore, $f_i(A) y_i = g_i(A) \phi(A) y_i = g_i(A) x_i = 0$. Therefore,

$$\sum_{j=1}^{\left|\beta_{y_j}\right|} a_{ij} A^{j-1} y_i = 0$$

and by independence of β_{y_i}, this implies $a_{ij} = 0$.

Next, it follows from the definition that for $W \equiv \operatorname{span}\left(\beta_{y_1}, \cdots, \beta_{y_p}\right)$,

$$\phi(A)(V) = \operatorname{span}\left(\beta_{x_1}, \cdots, \beta_{x_p}\right) \subseteq \phi(A) \operatorname{span}\left(\beta_{y_1}, \cdots, \beta_{y_p}\right) \equiv \phi(A)(W)$$

Now W is an A invariant subspace of $V = \ker(\phi(A)^m)$. Use Lemma 10.3.3 again to obtain $\beta_{z_1}, \cdots, \beta_{z_q}$ such that $\left\{\beta_{z_1}, \cdots, \beta_{z_q}, \beta_{y_1}, \cdots, \beta_{y_p}\right\}$ is a basis for $\ker(\phi(A)) + W$. From the above, $\phi(A)(W) = \phi(A)(V)$. Let $W' = W + \ker(\phi(A))$. Let U be the restriction of $\phi(A)$ to W' which is also $\phi(A)$ invariant. Then from the above inclusion, it follows that $U(W') = \phi(A)(V)$. Also $\ker(U) = \ker(\phi(A))$. This is because if $x \in \ker(\phi(A))$, then $x \in W'$ and so $Ux = 0$ also. If $Ux = 0$, then $x \in W'$ and $\phi(A) x = 0$. Hence $x \in \ker(\phi(A))$. Thus

$$\begin{aligned} \dim(W') &= \operatorname{rank}(U) + \dim(\ker(U)) \\ &= \operatorname{rank}(\phi(A)) + \dim(\ker(\phi(A))) = \dim(V) \end{aligned}$$

This shows $V = W'$ and so the above yields the desired basis. ■

10.4 Nilpotent Transformations

Definition 10.4.1 *Let V be a vector space over the field of scalars $\mathbb{F}$. Then $N \in \mathcal{L}(V, V)$ is called nilpotent if for some m, it follows that $N^m = 0$.*

The following lemma contains some significant observations about nilpotent transformations.

Lemma 10.4.2 *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 $\sum_{i=0}^{k} c_i N^i x = 0$. 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 $p(\lambda)$ is the minimal polynomial, then

$$p(\lambda) = \lambda^m l(\lambda) + r(\lambda)$$

where the degree of $r(\lambda)$ is less than m or else $r(\lambda) = 0$. Suppose the degree of $r(\lambda)$ is less than m. Then you would have

$$0 = 0 + r(N).$$

If $r(\lambda) = a_0 + a_1\lambda + \cdots + a_s\lambda^s$ for $s \leq m-1, a_s \neq 0$, then for any $x \in V$,

$$0 = a_0 x + a_1 N x + \cdots + a_s N^s x$$

If for some $x, N^s x \neq 0$, then from the first part of the argument, the above equation could not hold. Hence $N^s x = 0$ for all x and so $N^s = 0$ for some $s < m$, a contradiction to the choice of m. It follows that $r(\lambda) = 0$ and so $p(\lambda)$ cannot be the minimal polynomial unless $l(\lambda) = 1$. Hence $p(\lambda) = \lambda^m$ as claimed. ■

For such a nilpotent transformation, let $\left\{\beta_{x_1}, \cdots, \beta_{x_q}\right\}$ be a basis for $\ker(N^m) = V$ where these β_{x_i} are cyclic. This basis exists thanks to Theorem 10.3.4. Thus

$$V = \operatorname{span}\left(\beta_{x_1}\right) \oplus \cdots \oplus \operatorname{span}\left(\beta_{x_q}\right),$$

each of these subspaces in the above direct sum being N invariant. For x one of the x_k, consider β_x given by

$$x, Nx, N^2x, \cdots, N^{r-1}x$$

where $N^r x$ is in the span of the above vectors. Then by the above lemma, $N^r x = 0$.

By Theorem 10.2.5, 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 $\operatorname{span}\left(\beta_{x_k}\right)$. 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 $\beta_{x_{i+1}}$. 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 10.4.3 *$J_k(\alpha)$ 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 10.4.4 *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*

$$J = \begin{pmatrix} J_{r_1}(0) & & & 0 \\ & J_{r_2}(0) & & \\ & & \ddots & \\ 0 & & & J_{r_s}(0) \end{pmatrix}$$

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 a Jordan block of size* $r_j \times r_j$ *with 0 down the main diagonal.*

In fact, the matrix of the above proposition is unique.

Corollary 10.4.5 *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 10.4.4. 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

$$\begin{pmatrix} M_{r_1} & & & & & 0 \\ & \ddots & & & & \\ & & M_{r_k} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}, \begin{pmatrix} M_{r'_1} & & & & & 0 \\ & \ddots & & & & \\ & & M_{r'_k} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

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,

$$M_{r_k} = \begin{pmatrix} 0 & \cdots & 1 \\ & \ddots & \vdots \\ 0 & & 0 \end{pmatrix}$$

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. ■

10.5 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

$$p(\lambda) = \prod_{k=1}^{r} (\lambda - \lambda_k)^{m_k}$$

Recall the following which follows from Theorem 9.4.4.

Proposition 10.5.1 *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 10.2.3 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 10.4.4, 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

$$\begin{pmatrix} J(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_r) \end{pmatrix} \tag{10.6}$$

where $J(\lambda_k)$ is an $m_k \times m_k$ matrix of the form

$$\begin{pmatrix} J_{k_1}(\lambda_k) & & & 0 \\ & J_{k_2}(\lambda_k) & & \\ & & \ddots & \\ 0 & & & J_{k_r}(\lambda_k) \end{pmatrix} \tag{10.7}$$

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

$$\begin{pmatrix} \lambda & 1 & & 0 \\ 0 & \lambda & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 0 & \lambda \end{pmatrix} \tag{10.8}$$

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 10.5.2 *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 10.6 - 10.8, 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, $\operatorname{rank}\left(A^k\right) = \operatorname{rank}\left(B^k\right)$ 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 $\{\operatorname{rank}((J(\lambda_k) - \lambda_k I)^m)\}_{m=1}^{\infty}$ and $\left\{\operatorname{rank}\left((J'(\lambda_k) - \lambda_k I)^m\right)\right\}_{m=1}^{\infty}$ 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 10.4.5. ■

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

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

Therefore, $A^2 = S^{-1} J S S^{-1} J S = S^{-1} J^2 S$ and continuing this way, it follows

$$A^k = S^{-1} J^k S.$$

[1]The S here is written as S^{-1} in the corollary.

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

$$J_s = \begin{pmatrix} \alpha & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha \end{pmatrix} \tag{10.9}$$

which can be written in the form

$$J_s = D + N$$

for D a multiple of the identity and N an upper triangular matrix with zeros down the main diagonal. Therefore, by the Cayley Hamilton theorem, $N^{m_s} = 0$ because the characteristic equation for N is just $\lambda^{m_s} = 0$. (You could also verify this directly.) 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 $k \geq m_s$.

$$\begin{aligned} J_s^k &= (D+N)^k = \sum_{j=0}^{k} \binom{k}{j} D^{k-j} N^j \\ &= \sum_{j=0}^{m_s} \binom{k}{j} D^{k-j} N^j, \end{aligned} \tag{10.10}$$

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 10.5.3 *Suppose J is of the form J_s described above in 10.9 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 10.10, 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. ■

10.6 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

$$p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k = 0,$$

   when the value of the polynomial equals zero whenever a particular value of $\lambda \in \mathbb{F}$ is placed in the formula for $p(\lambda)$, can the same conclusion be drawn? Is there any difference in the two definitions for ordinary fields like $\mathbb{Q}$? **Hint:** Consider $\mathbb{Z}_2$, the integers mod 2.

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 8.3.11 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 $\left\{\mathbf{v}, A\mathbf{v}, \cdots, A^{m-1}\mathbf{v}\right\}$ 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
$$\det(\Psi)' = \text{trace}(A)\det(\Psi)$$
and so
$$\det(\Psi(t)) = Ce^{\text{trace}(A)t}$$
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
$$\det(\Phi)' = \sum_{i=1}^{n} \det(\Phi_i) \tag{10.11}$$
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 10.11 and use properties of determinants to obtain
$$\det(\Phi)' = \sum_{i=1}^{n} \lambda_i \det(\Phi).$$
Obtain Abel's formula
$$\det(\Phi) = Ce^{\text{trace}(A)t}$$
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. In proving the uniqueness of the Jordan canonical form, it was asserted 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 10.3.4 show that the cyclic sets can be arranged in such a way that the length of $\beta_{v_{i+1}}$ divides the length of β_{v_i}.

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 $\ker (\phi_i (A)^{m_i})$. 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 $q (A) x = 0$.

10.7 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 [9] and there are more details given in this reference. Another good source which has additional results is [14].

Here is a definition of the concept of a companion matrix.

Definition 10.7.1 *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 10.7.2 *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

$$\mathbf{e}_k = C^{k-1}\mathbf{e}_1,\ k = 1, \cdots, n \tag{10.12}$$

and so it follows

$$\left\{\mathbf{e}_1, C\mathbf{e}_1, C^2\mathbf{e}_1, \cdots, C^{n-1}\mathbf{e}_1\right\} \tag{10.13}$$

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 10.12 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 10.13 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 10.7.3 *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* $V_k \equiv \ker(\phi_k(\lambda)^{m_k})$, *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

$$\begin{pmatrix} M^1 & & 0 \\ & \ddots & \\ 0 & & M^q \end{pmatrix} \tag{10.14}$$

If B_k is given as $\left\{\beta_{v_1}, \cdots, \beta_{v_s}\right\}$ as described in Theorem 10.3.4 where each β_{v_j} is an A cyclic set of vectors, then the matrix M^k is of the form

$$M^k = \begin{pmatrix} C\left(\phi_k(\lambda)^{r_1}\right) & & 0 \\ & \ddots & \\ 0 & & C\left(\phi_k(\lambda)^{r_s}\right) \end{pmatrix} \tag{10.15}$$

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 10.2.5 the matrix of A with respect to $\{B_1, \cdots, B_q\}$ is of the form given in 10.14. Now by Theorem 10.3.4 the basis B_k may be chosen in the form $\left\{\beta_{v_1}, \cdots, \beta_{v_s}\right\}$ 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 $\text{span}\left(\beta_{v_k}\right)$. Then you can apply Theorem 10.2.5 to get the result in 10.15. 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 by Corollary 8.3.11, $\eta(\lambda) = \phi_k(\lambda)^{r_k}$ where $r_k \leq m_k$. It remains to consider the matrix of A restricted to $\text{span}\left(\beta_{v_k}\right)$. Say

$$\eta(\lambda) = \phi_k(\lambda)^{r_k} = a_0 + a_1\lambda + \cdots + a_{d-1}\lambda^{d-1} + \lambda^d$$

Thus

$$A^d v_k = -a_0 v_k - a_1 A v_k - \cdots - a_{d-1} A^{d-1} v_k$$

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. The 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. ■

10.8 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 10.8.1 *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 $A^m x \in \operatorname{span}\left(x, \cdots, A^{m-1}x\right)$.*

The following proposition ties these A cyclic sets to polynomials. It is just a review of ideas used above to prove existence.

Proposition 10.8.2 *Let $x \neq 0$ and consider $\left\{x, Ax, A^2x, \cdots, A^{m-1}x\right\}$. 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 $p \leq m, \eta(\lambda) = \phi(\lambda)^p$.*

Lemma 10.8.3 *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 $\left\{\beta_{v_1}, \cdots, \beta_{v_s}\right\}$ 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 $\left|\beta_{v_k}\right|$ denote the number of vectors in β_{v_k}, it follows there is only one possible set of numbers $\left|\beta_{v_k}\right|$.*

Proof: Say β_{v_j} is associated with the polynomial $\phi(\lambda)^{p_j}$. Thus, as described above $\left|\beta_{v_j}\right|$ equals $p_j d$. Consider the following table which comes from the A cyclic set

$$\left\{v_j, Av_j, \cdots, A^{d-1}v_j, \cdots, A^{p_j d-1}v_j\right\}$$

α_0^j	α_1^j	α_2^j	$\cdots$	α_{d-1}^j
v_j	Av_j	A^2v_j	$\cdots$	$A^{d-1}v_j$
$\phi(A)v_j$	$\phi(A)Av_j$	$\phi(A)A^2v_j$	$\cdots$	$\phi(A)A^{d-1}v_j$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$\phi(A)^{p_j-1}v_j$	$\phi(A)^{p_j-1}Av_j$	$\phi(A)^{p_j-1}A^2v_j$	$\cdots$	$\phi(A)^{p_j-1}A^{d-1}v_j$

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

$$\sum_{i=0}^{d-1} c_{i0} \phi(A)^{p_j-1} A^i v_j = 0$$

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. 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 10.4.4 and Theorem 10.5.2 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 10.8.4 *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 10.2.3

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

where $V_k = \ker\left(\phi_k(A)^{m_k}\right)$. Also recall from Corollary 10.2.4 that the minimal polynomial of the restriction of A to V_k is $\phi_k(\lambda)^{m_k}$. Now apply Lemma 10.8.3 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 10.8.5 *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 characteristic polynomial is $\lambda^3 - 24\lambda^2 + 180\lambda - 432$. This factors as

$$(\lambda-6)^2(\lambda-12)$$

It turns out this is also the minimal polynomial. You can see this by plugging in A where you see λ and observing things don't work if you delete one of the $\lambda - 6$ factors. There is more on this in the exercises. It turns out you can compute the minimal polynomial pretty easily. 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 9.3.10, 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 10.8.6 *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.

10.9 Exercises

1. Letting A be a complex $n \times n$ matrix, in obtaining the rational canonical form, one obtains $\mathbb{C}^n$ as a direct sum of the form
$$\text{span}\left(\beta_{\mathbf{x}_1}\right) \oplus \cdots \oplus \text{span}\left(\beta_{\mathbf{x}_r}\right)$$
where β_x is an ordered cyclic set of vectors, $\mathbf{x}, A\mathbf{x}, \cdots, A^{m-1}\mathbf{x}$ such that $A^m\mathbf{x}$ is in the span of the previous vectors. Now apply the Gram Schmidt process to the ordered basis $\left(\beta_{\mathbf{x}_1}, \beta_{\mathbf{x}_2}, \cdots, \beta_{\mathbf{x}_r}\right)$, the vectors in each $\beta_{\mathbf{x}_i}$ listed according to increasing power of A, thus obtaining an ordered basis $(\mathbf{q}_1, \cdots, \mathbf{q}_n)$. Letting Q be the unitary matrix which has these vectors as columns, show that Q^*AQ equals a matrix B which satisfies $B_{ij} = 0$ if $i - j \geq 2$. Such a matrix is called an upper Hessenberg matrix and this shows that every $n \times n$ matrix is orthogonally similar to an upper Hessenberg matrix. These are zero below the main sub diagonal, like companion matrices discussed above.

2. In the argument for Theorem 10.2.3 it was shown that $m(A)\phi_l(A)v = v$ whenever $v \in \ker\left(\phi_k(A)^{r_k}\right)$. Show that $m(A)$ restricted to $\ker\left(\phi_k(A)^{r_k}\right)$ is the inverse of the linear transformation $\phi_l(A)$ on $\ker\left(\phi_k(A)^{r_k}\right)$.

3. 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 8.3.11 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.

4. 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?

5. 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.

6. Letting the field of scalars be $\mathbb{Q}$, find the rational canonical form for the matrix
$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 2 & 3 & 0 & 2 \\ 1 & 3 & 2 & 4 \\ 1 & 2 & 1 & 2 \end{pmatrix}$$

7. 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?

8. 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 $\mathbb{Q} + i\mathbb{Q}$.

9. 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 $q(\lambda)$.

10. $\uparrow$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.

11. 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 10.8.3 and show that they have the same rational canonical form with respect to $\mathbb{F}$.

Markov Processes

11.1 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 11.1.1 *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. A matrix which has nonnegative entries such that

$$\sum_j a_{ij} = 1$$

will also be called a stochastic 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 11.1.2 *The property of being a stochastic matrix is preserved by taking products.*

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,

$$\sum_i \sum_k a_{ik} b_{kj} = \sum_k \left(\sum_i a_{ik} \right) b_{kj} = \sum_k b_{kj} = 1.$$

A similar argument yields the same result in the case where it is the sum over a column which is equal to 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.

The following theorem is convenient for showing the existence of limits.

Theorem 11.1.3 *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 11.1.2. 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

$$P^{-1}A^nP = J^n \tag{11.1}$$

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 11.1.4 *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 $\mu = 1$. In addition to this, if $A\mathbf{v} = \mathbf{v}$ for a nonzero vector, $\mathbf{v} \in \mathbb{R}^n$, then $v_j v_i \geq 0$ for all i, j so the components of $\mathbf{v}$ have the same sign.*

Proof: Suppose the matrix satisfies

$$\sum_j a_{ij} = 1.$$

Then if $\mathbf{v} = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}^T$, it is obvious that $A\mathbf{v} = \mathbf{v}$. Therefore, this matrix has $\lambda = 1$ as an eigenvalue. Suppose then that μ is an eigenvalue. Is $|\mu| < 1$ or $\mu = 1$? Let $\mathbf{v}$ be an eigenvector and let $|v_i|$ be the largest of the $|v_j|$.

$$\mu v_i = \sum_j a_{ij} v_j$$

and now multiply both sides by $\overline{\mu v_i}$ to obtain

$$\begin{aligned} |\mu|^2 |v_i|^2 &= \sum_j a_{ij} v_j \overline{v_i \mu} = \sum_j a_{ij} \operatorname{Re}\left(v_j \overline{v_i \mu}\right) \\ &\leq \sum_j a_{ij} |\mu| |v_i|^2 = |\mu| |v_i|^2 \end{aligned}$$

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 $j = 1$ which shows $\overline{\mu}$ and hence μ are real. Thus, in this case, $\mu = 1$. The only other case is where $|\mu| < 1$.

If instead, $\sum_i a_{ij} = 1$, consider A^T. Both A and A^T have the same characteristic polynomial and so their eigenvalues are exactly the same. ■

Lemma 11.1.5 *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

$$\sum_i w_i = \sum_i \sum_j a_{ij} v_j = \sum_j \sum_i a_{ij} v_j = \sum_j v_j = c.$$

The following theorem about limits is now easy to obtain.

Theorem 11.1.6 *Suppose A is a Markov matrix (The sum over a column equals 1) 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 11.1.4, since each $a_{ij} > 0$, the eigenvalues are either 1 or have absolute value less than 1. Therefore, the claimed limit exists by Theorem 11.1.3. The assertion that the components are nonnegative and sum to c follows from Lemma 11.1.5. 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 11.1.7 *Suppose A is a regular Markov matrix, on 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. Now suppose that $a_{ij} \geq 0$ for all i, j and $A = (a_{ij})$ is a transition matrix. Then if $B = (b_{ij})$ is a transition matrix with $b_{ij} > 0$ for all ij, it follows that BA is a transition matrix which has strictly positive entries. The ij^{th} entry of BA is

$$\sum_k b_{ik} a_{kj} > 0,$$

Thus, from Lemma 11.1.4, 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, for if λ is an eigenvalue of A with $|\lambda| = 1$, then λ^k is an eigenvalue for A^k and so, for all k large enough, $\lambda^k = 1$ which is absurd unless $\lambda = 1$. By Theorem 11.1.3, $\lim_{n\to\infty} A^n \mathbf{w}$ exists. The rest follows as in Theorem 11.1.6. $\blacksquare$

11.2 Migration Matrices

Definition 11.2.1 *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 11.1.7 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

$$\begin{aligned} y &= x \\ z &= 4x \\ x + y + z &= 150000 \end{aligned}$$

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.

11.3 Markov Chains

A random variable is just a function which can have certain values which have probabilities associated with them. Thus it makes sense to consider the probability that the random variable has a certain value or is in some set. The idea of a Markov chain is a sequence of random variables, $\{X_n\}$ which can be in any of a collection of states which can be labeled with nonnegative integers. Thus you can speak of the probability the random variable, X_n is in state i. The probability that X_{n+1} is in state j given that X_n is in state i is called a one step transition probability. When this probability does not depend on n it is called stationary and this is the case of interest here. Since this probability does not depend on n it can be denoted by p_{ij}. Here is a simple example called a random walk.

Example 11.3.1 *Let there be n points, x_i, and consider a process of something moving randomly from one point to another. Suppose X_n is a sequence of random variables which has values $\{1, 2, \cdots, n\}$ where $X_n = i$ indicates the process has arrived at the i^{th} point. Let p_{ij} be the probability that X_{n+1} has the value j given that X_n has the value i. Since X_{n+1} must have some value, it must be the case that $\sum_j a_{ij} = 1$. Note this says that the sum over a row equals 1 and so the situation is a little different than the above in which the sum was over a column.*

As an example, let x_1, x_2, x_3, x_4 be four points taken in order on $\mathbb{R}$ and suppose x_1 and x_4 are absorbing. This means that $p_{4k} = 0$ for all $k \neq 4$ and $p_{1k} = 0$ for all $k \neq 1$. Otherwise, you can move either to the left or to the right with probability $\frac{1}{2}$. The Markov matrix associated with this situation is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Definition 11.3.2 *Let the stationary transition probabilities, p_{ij} be defined above. The resulting matrix having p_{ij} as its ij^{th} entry is called the matrix of transition probabilities. The sequence of random variables for which these p_{ij} are the transition probabilities is called a Markov chain. The matrix of transition probabilities is called a stochastic matrix.*

The next proposition is fundamental and shows the significance of the powers of the matrix of transition probabilities.

Proposition 11.3.3 *Let p_{ij}^n denote the probability that X_n is in state j given that X_0 was in state i. Then p_{ij}^n is the ij^{th} entry of the matrix P^n where $P = (p_{ij})$.*

Proof: This is clearly true if $n = 1$ and follows from the definition of the p_{ij}. Suppose true for n. Then the probability that X_{n+1} is at j given that X_0 was at i equals $\sum_k p_{ik}^n p_{kj}$ because X_n must have some value, k, and so this represents all possible ways to go from i to j. You can go from i to 1 in n steps with probability p_{i1} and then from 1 to j in one step with probability p_{1j} and so the probability of this is $p_{i1}^n p_{1j}$ but you can also go from i to 2 and then from 2 to j and from i to 3 and then from 3 to j etc. Thus the sum of these is just what is given and represents the probability of X_{n+1} having the value j given X_0 has the value i. ■

In the above random walk example, lets take a power of the transition probability matrix to determine what happens. Rounding off to two decimal places,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{20} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ .67 & 9.5 \times 10^{-7} & 0 & .33 \\ .33 & 0 & 9.5 \times 10^{-7} & .67 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus p_{21} is about $2/3$ while p_{32} is about $1/3$ and terms like p_{22} are very small. You see this seems to be converging to the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

After many iterations of the process, if you start at 2 you will end up at 1 with probability 2/3 and at 4 with probability 1/3. This makes good intuitive sense because it is twice as far from 2 to 4 as it is from 2 to 1.

Theorem 11.3.4 *The eigenvalues of*

$$\begin{pmatrix} 0 & p & 0 & \cdots & 0 \\ q & 0 & p & \cdots & 0 \\ 0 & q & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & p \\ 0 & \vdots & 0 & q & 0 \end{pmatrix}$$

have absolute value less than 1. Here $p + q = 1$ *and both* $p, q > 0$.

Proof: By Gerschgorin's theorem, if λ is an eigenvalue, then $|\lambda| \leq 1$. Now suppose $\mathbf{v}$ is an eigenvector for λ. Then

$$A\mathbf{v} = \begin{pmatrix} pv_2 \\ qv_1 + pv_3 \\ \vdots \\ qv_{n-2} + pv_n \\ qv_{n-1} \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}$$

Suppose $|\lambda| = 1$. Then the top row shows $p\,|v_2| = |v_1|$ so $|v_1| < |v_2|$. Suppose $|v_1| < |v_2| < \cdots < |v_k|$ for some $k < n$. Then

$$|\lambda v_k| = |v_k| \leq q\,|v_{k-1}| + p\,|v_{k+1}| < q\,|v_k| + p\,|v_{k+1}|$$

and so subtracting $q\,|v_k|$ from both sides,

$$p\,|v_k| < p\,|v_{k+1}|$$

showing $\{|v_k|\}_{k=1}^{n}$ is an increasing sequence. Now a contradiction results on the last line which requires $|v_{n-1}| > |v_n|$. Therefore, $|\lambda| < 1$ for any eigenvalue of the above matrix. ■

Corollary 11.3.5 *Let p, q be positive numbers and let $p + q = 1$. The eigenvalues of*

$$\begin{pmatrix} a & p & 0 & \cdots & 0 \\ q & a & p & \cdots & 0 \\ 0 & q & a & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & p \\ 0 & \vdots & 0 & q & a \end{pmatrix}$$

are all strictly closer than 1 to a. That is, whenever λ is an eigenvalue,

$$|\lambda - a| < 1$$

have absolute value less than 1.

Proof: Let A be the above matrix and suppose $A\mathbf{x} = \lambda\mathbf{x}$. Then letting A' denote

$$\begin{pmatrix} 0 & p & 0 & \cdots & 0 \\ q & 0 & p & \cdots & 0 \\ 0 & q & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & p \\ 0 & \vdots & 0 & q & 0 \end{pmatrix},$$

it follows

$$A'\mathbf{x} = (\lambda - a)\,\mathbf{x}$$

and so from the above theorem,

$$|\lambda - a| < 1. ■$$

Example 11.3.6 *In the gambler's ruin problem a gambler plays a game with someone, say a casino, until he either wins all the other's money or loses all of his own. A simple version of this is as follows. Let X_k denote the amount of money the gambler has. Each time the game is played he wins with probability $p \in (0, 1)$ or loses with probability $(1 - p) \equiv q$. In case he wins, his money increases to $X_k + 1$ and if he loses, his money decreases to $X_k - 1$.*

The transition probability matrix P, describing this situation is as follows.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & \vdots \\ 0 & 0 & q & 0 & \ddots & \vdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & p & 0 \\ 0 & 0 & \vdots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{11.2}$$

Here the matrix is $b+1 \times b+1$ because the possible values of X_k are all integers from 0 up to b. The 1 in the upper left corner corresponds to the gambler's ruin. It involves $X_k = 0$ so he has no money left. Once this state has been reached, it is not possible to ever leave it. This is indicated by the row of zeros to the right of this entry the k^{th} of which gives the probability of going from state 1 corresponding to no money to state k[1].

In this case 1 is a repeated root of the characteristic equation of multiplicity 2 and all the other eigenvalues have absolute value less than 1. To see that this is the case, note that the characteristic polynomial is of the form

$$(1-\lambda)^2 \det \begin{pmatrix} -\lambda & p & 0 & \cdots & 0 \\ q & -\lambda & p & \cdots & 0 \\ 0 & q & -\lambda & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & p \\ 0 & \vdots & 0 & q & -\lambda \end{pmatrix}$$

and the factor after $(1-\lambda)^2$ has zeros which are in absolute value less than 1. Its zeros are the eigenvalues of the matrix

$$A \equiv \begin{pmatrix} 0 & p & 0 & \cdots & 0 \\ q & 0 & p & \cdots & 0 \\ 0 & q & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & p \\ 0 & \vdots & 0 & q & 0 \end{pmatrix}$$

and by Corollary 11.3.5 these all have absolute value less than 1.

Therefore, by Theorem 11.1.3 $\lim_{n\to\infty} P^n$ exists. The case of $\lim_{n\to\infty} p^n_{j0}$ is particularly interesting because it gives the probability that, starting with an amount j, the gambler eventually ends up at 0 and is **ruined**. From the matrix, it follows

$$\begin{aligned} p^n_{j0} &= qp^{n-1}_{(j-1)0} + pp^{n-1}_{(j+1)0} \text{ for } j \in [1, b-1], \\ p^n_{00} &= 1, \text{ and } p^n_{b0} = 0. \end{aligned}$$

To simplify the notation, define $P_j \equiv \lim_{n\to\infty} p^n_{j0}$ as the probability of ruin given the initial fortune of the gambler equals j. Then the above simplifies to

$$\begin{aligned} P_j &= qP_{j-1} + pP_{j+1} \text{ for } j \in [1, b-1], \\ P_0 &= 1, \text{ and } P_b = 0. \end{aligned} \tag{11.3}$$

[1] No one will give the gambler money. This is why the only reasonable number for entries in this row to the right of 1 is 0.

Now, knowing that P_j exists, it is not too hard to find it from 11.3. This equation is called a difference equation and there is a standard procedure for finding solutions of these. You try a solution of the form $P_j = x^j$ and then try to find x such that things work out. Therefore, substitute this in to the first equation of 11.3 and obtain

$$x^j = qx^{j-1} + px^{j+1}.$$

Therefore,

$$px^2 - x + q = 0$$

and so in case $p \neq q$, you can use the fact that $p + q = 1$ to obtain

$$\begin{aligned} x &= \frac{1}{2p}\left(1+\sqrt{(1-4pq)}\right) \text{ or } \frac{1}{2p}\left(1-\sqrt{(1-4pq)}\right) \\ &= \frac{1}{2p}\left(1+\sqrt{(1-4p(1-p))}\right) \text{ or } \frac{1}{2p}\left(1-\sqrt{(1-4p(1-p))}\right) \\ &= 1 \text{ or } \frac{q}{p}. \end{aligned}$$

Now it follows that both $P_j = 1$ and $P_j = \left(\frac{q}{p}\right)^j$ satisfy the difference equation 11.3. Therefore, anything of the form

$$\alpha + \beta \left(\frac{q}{p}\right)^j \tag{11.4}$$

will satisfy this equation. Find a, b such that this also satisfies the second equation of 11.3. Thus it is required that

$$\alpha + \beta = 1, \ \alpha + \beta \left(\frac{q}{p}\right)^b = 0$$

and so

$$\alpha + \beta = 1, \ \alpha + \beta \left(\frac{q}{p}\right)^b = 0$$

Solution is : $\left\{\beta = -\frac{1}{-1+\left(\frac{q}{p}\right)^b}, \alpha = \frac{\left(\frac{q}{p}\right)^b}{-1+\left(\frac{q}{p}\right)^b}\right\}$. Substituting this in to 11.4 and simplifying, yields the following in the case that $p \neq q$.

$$P_j = \frac{p^{b-j}q^j - q^b}{p^b - q^b} \tag{11.5}$$

Note that

$$\lim_{p \to q} \frac{p^{b-j}q^j - q^b}{p^b - q^b} = \frac{b-j}{b}.$$

Thus as the game becomes more fair in the sense the probabilities of winning become closer to $1/2$, the probability of ruin given an initial amount j is $\frac{b-j}{b}$.

Alternatively, you could consider the difference equation directly in the case where $p = q = 1/2$. In this case, you can see that two solutions to the difference equation

$$\begin{aligned} P_j &= \frac{1}{2}P_{j-1} + \frac{1}{2}P_{j+1} \text{ for } j \in [1, b-1], \\ P_0 &= 1, \text{ and } P_b = 0. \end{aligned} \tag{11.6}$$

are $P_j = 1$ and $P_j = j$. This leads to a solution to the above of

$$P_j = \frac{b-j}{b}. \tag{11.7}$$

This last case is pretty interesting because it shows, for example that if the gambler starts with a fortune of 1 so that he starts at state $j = 1$, then his probability of losing all is $\frac{b-1}{b}$ which might be quite large, especially if the other player has a lot of money to begin with. As the gambler starts with more and more money, his probability of losing everything does decrease.

11.4 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 11.1.3 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. $\uparrow$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. $\uparrow$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 $\begin{pmatrix} \alpha^2 & \alpha\mathbf{b}^* \\ \alpha\mathbf{b} & A \end{pmatrix}, \alpha$ real.

19. $\uparrow$ Suppose A is a nonnegative Hermitian matrix 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. $\uparrow$ Prove the inequality of Hadamard. If A is a Hermitian matrix which is nonnegative, then

$$\det(A) \leq \prod_i A_{ii}$$

Inner Product Spaces

12.1 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 12.1.1 *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 12.1.2 *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 12.1.3 *A normed linear space in which the norm comes from an inner product as just described is called an inner product space.*

Example 12.1.4 *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 12.1.5 *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 12.1.6 *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 12.1.7 *(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 $\overline{\omega}(x, y) = |(x, y)| = \operatorname{Re}(x, y\omega)$. Let

$$F(t) = (x + ty\omega, x + t\omega y).$$

Then from the axioms of the inner product,

$$F(t) = |x|^2 + 2t \operatorname{Re}(x, \omega y) + t^2 |y|^2 \geq 0.$$

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 12.1.8 *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,

$$\begin{aligned} |x+y|^2 &\equiv (x+y,x+y) \equiv |x|^2 + |y|^2 + 2\operatorname{Re}(x,y) \\ &\leq |x|^2 + |y|^2 + 2|(x,y)| \\ &\leq |x|^2 + |y|^2 + 2|x||y| = (|x|+|y|)^2. \blacksquare \end{aligned}$$

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 12.1.9 *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 12.1.10 *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,

$$0 = \sum_{j=1}^{n} c^j (u_j, u_k) = \sum_{j=1}^{n} c^j \delta_{jk} = c^k.$$

12.2 The Gram Schmidt Process

Lemma 12.2.1 *Let* X *be a finite dimensional inner product space of dimension* n *whose basis is* $\{x_1, \cdots, x_n\}$*. Then there exists an orthonormal basis for* X*,* $\{u_1, \cdots, u_n\}$ *which has the property that for each* $k \leq n$*, span*$(x_1, \cdots, x_k) = \operatorname{span}(u_1, \cdots, u_k)$*.*

Proof: Let $\{x_1, \cdots, x_n\}$ be a basis for X. Let $u_1 \equiv x_1/|x_1|$. Thus for $k=1$, $\operatorname{span}(u_1) = \operatorname{span}(x_1)$ 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 $\operatorname{span}(x_1, \cdots, x_k) = \operatorname{span}(u_1, \cdots, u_k)$. Then define

$$u_{k+1} \equiv \frac{x_{k+1} - \sum_{j=1}^{k} (x_{k+1}, u_j) u_j}{\left| x_{k+1} - \sum_{j=1}^{k} (x_{k+1}, u_j) u_j \right|}, \tag{12.1}$$

where the denominator is not equal to zero because the x_j form a basis and so

$$x_{k+1} \notin \operatorname{span}(x_1, \cdots, x_k) = \operatorname{span}(u_1, \cdots, u_k)$$

Thus by induction,

$$u_{k+1} \in \operatorname{span}(u_1, \cdots, u_k, x_{k+1}) = \operatorname{span}(x_1, \cdots, x_k, x_{k+1}).$$

Also, $x_{k+1} \in \operatorname{span}(u_1, \cdots, u_k, u_{k+1})$ which is seen easily by solving 12.1 for x_{k+1} and it follows

$$\operatorname{span}(x_1, \cdots, x_k, x_{k+1}) = \operatorname{span}(u_1, \cdots, u_k, u_{k+1}).$$

If $l \leq k$,

$$\begin{aligned}
(u_{k+1}, u_l) &= C\left((x_{k+1}, u_l) - \sum_{j=1}^{k} (x_{k+1}, u_j)(u_j, u_l)\right) \\
&= C\left((x_{k+1}, u_l) - \sum_{j=1}^{k} (x_{k+1}, u_j)\,\delta_{lj}\right) \\
&= C\left((x_{k+1}, u_l) - (x_{k+1}, u_l)\right) = 0.
\end{aligned}$$

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 12.2.2 *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 12.2.3 *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: By assumption that this is an orthonormal basis,

$$\sum_{j=1}^{n} (x, u_j) \overbrace{(u_j, u_l)}^{\delta_{jl}} = (x, u_l) .$$

Letting $y = \sum_{k=1}^{n} (x, u_k) u_k$, it follows

$$\begin{aligned} (x - y, u_j) &= (x, u_j) - \sum_{k=1}^{n} (x, u_k) (u_k, u_j) \\ &= (x, u_j) - (x, u_j) = 0 \end{aligned}$$

for all j. Hence, for any choice of scalars $c^1, \cdots, c^n$,

$$\left(x - y, \sum_{j=1}^{n} c^j u_j \right) = 0$$

and so $(x - y, z) = 0$ for all $z \in X$. Thus this holds in particular for $z = x - y$. Therefore, $x = y$. ■

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 12.2.4 *Let M be a subspace of X, a finite dimensional inner product space and let $\{x_i\}_{i=1}^m$ be an orthonormal basis for M. Then if $y \in X$ and $w \in M$,*

$$|y - w|^2 = \inf \left\{ |y - z|^2 : z \in M \right\} \tag{12.2}$$

if and only if

$$(y - w, z) = 0 \tag{12.3}$$

for all $z \in M$. Furthermore,

$$w = \sum_{i=1}^{m} (y, x_i) x_i \tag{12.4}$$

is the unique element of M which has this property. It is called the orthogonal projection.

Proof: Let $t \in \mathbb{R}$. Then from the properties of the inner product,

$$|y - (w + t(z - w))|^2 = |y - w|^2 + 2t \operatorname{Re}(y - w, w - z) + t^2 |z - w|^2 . \tag{12.5}$$

If $(y - w, z) = 0$ for all $z \in M$, then letting $t = 1$, the middle term in the above expression vanishes and so $|y - z|^2$ is minimized when $z = w$.

Conversely, if 12.2 holds, then the middle term of 12.5 must also vanish since otherwise, you could choose small real t such that

$$|y - w|^2 > |y - (w + t(z - w))|^2 .$$

Here is why. If $\operatorname{Re}(y - w, w - z) < 0$, then let t be very small and positive. The middle term in 12.5 will then be more negative than the last term is positive and the right side of this formula will then be less than $|y - w|^2$. If $\operatorname{Re}(y - w, w - z) > 0$ then choose t small and negative to achieve the same result.

It follows, letting $z_1 = w - z$ that

$$\operatorname{Re}(y - w, z_1) = 0$$

for all $z_1 \in M$. Now letting $\omega \in \mathbb{C}$ be such that $\omega (y - w, z_1) = |(y - w, z_1)|$,

$$|(y - w, z_1)| = (y - w, \overline{\omega} z_1) = \operatorname{Re}(y - w, \overline{\omega} z_1) = 0,$$

which proves the first part of the theorem since z_1 is arbitrary.

It only remains to verify that w given in 12.4 satisfies 12.3 and is the only point of M which does so. To do this, note that if c_i, d_i are scalars, then the properties of the inner product and the fact the $\{x_i\}$ are orthonormal implies

$$\left(\sum_{i=1}^{m} c_i x_i, \sum_{j=1}^{m} d_j x_j \right) = \sum_i c_i \overline{d_i}.$$

By Lemma 12.2.3,

$$z = \sum_i (z, x_i) x_i$$

and so

$$\begin{aligned}
\left(y - \sum_{i=1}^{m} (y, x_i) x_i, z \right) &= \left(y - \sum_{i=1}^{m} (y, x_i) x_i, \sum_{i=1}^{m} (z, x_i) x_i \right) \\
&= \sum_{i=1}^{m} \overline{(z, x_i)} (y, x_i) - \left(\sum_{i=1}^{m} (y, x_i) x_i, \sum_{j=1}^{m} (z, x_j) x_j \right) \\
&= \sum_{i=1}^{m} \overline{(z, x_i)} (y, x_i) - \sum_{i=1}^{m} (y, x_i) \overline{(z, x_i)} = 0.
\end{aligned}$$

This shows w given in 12.4 does minimize the function, $z \to |y - z|^2$ for $z \in M$. It only remains to verify uniqueness. Suppose than that $w_i, i = 1, 2$ minimizes this function of z for $z \in M$. Then from what was shown above,

$$\begin{aligned}
|y - w_1|^2 &= |y - w_2 + w_2 - w_1|^2 \\
&= |y - w_2|^2 + 2 \operatorname{Re}(y - w_2, w_2 - w_1) + |w_2 - w_1|^2 \\
&= |y - w_2|^2 + |w_2 - w_1|^2 \le |y - w_2|^2 ,
\end{aligned}$$

the last equal sign holding because w_2 is a minimizer and the last inequality holding because w_1 minimizes. ■

12.3 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 12.3.1 *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 12.2.1, there exists an orthonormal basis, $\{u_j\}_{j=1}^n$. Define

$$z \equiv \sum_{j=1}^n \overline{f(u_j)} u_j.$$

Then using Lemma 12.2.3,

$$\begin{aligned}(x, z) &= \left(x, \sum_{j=1}^n \overline{f(u_j)} u_j\right) = \sum_{j=1}^n f(u_j)(x, u_j) \\ &= f\left(\sum_{j=1}^n (x, u_j) u_j\right) = f(x). \blacksquare\end{aligned}$$

Corollary 12.3.2 *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*

$$(Ax, y)_Y = (x, A^* y)_X \tag{12.6}$$

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 $x \in X$,

$$\begin{gathered}(x, A^*(ay_1 + by_2))_X \equiv (Ax, (ay_1 + by_2))_Y \\ \equiv \overline{a}(Ax, y_1) + \overline{b}(Ax, y_2) \equiv \\ \overline{a}(x, A^*(y_1)) + \overline{b}(x, A^*(y_2)) = (x, aA^*(y_1) + bA^*(y_2)) .\end{gathered}$$

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.

$$\left(x, (\alpha A + \beta B)^* y\right) \equiv ((\alpha A + \beta B) x, y)$$

$$\begin{aligned} &= \alpha (Ax, y) + \beta (Bx, y) = \alpha (x, A^* y) + \beta (x, B^* y) \\ &= (x, \overline{\alpha} A^* y) + \left(x, \overline{\beta} A^* y\right) = \left(x, \left(\overline{\alpha} A^* + \overline{\beta} A^*\right) y\right). \end{aligned}$$

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 12.3.3 *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 12.3.4 *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,

$$M_{ij}^* = \overline{M_{ji}}$$

which gives the conclusion of the theorem.

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 12.3.5 *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*

$$QV \subseteq \operatorname{span}(\mathbf{e}_1, \cdots, \mathbf{e}_p)$$

and $|Q\mathbf{x}| = |\mathbf{x}|$ for all $\mathbf{x}$. Also

$$Q^* Q = QQ^* = I.$$

Proof: By Lemma 12.2.1 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

$$(Q(\mathbf{x}+\mathbf{y}), Q(\mathbf{x}+\mathbf{y})) = (\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}).$$

Thus

$$|Q\mathbf{x}|^2 + |Q\mathbf{y}|^2 + 2\operatorname{Re}(Q\mathbf{x}, Q\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\operatorname{Re}(\mathbf{x}, \mathbf{y})$$

and since Q preserves norms, it follows that for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$,

$$\operatorname{Re}(Q\mathbf{x}, Q\mathbf{y}) = \operatorname{Re}(\mathbf{x}, Q^*Q\mathbf{y}) = \operatorname{Re}(\mathbf{x}, \mathbf{y}).$$

Thus

$$\operatorname{Re}(\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y}) = 0 \tag{12.7}$$

for all $\mathbf{x}, \mathbf{y}$. Let ω be a complex number such that $|\omega| = 1$ and

$$\omega(\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y}) = |(\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y})|.$$

Then from 12.7,

$$\begin{aligned} 0 &= \operatorname{Re}(\omega\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y}) = \operatorname{Re}\omega(\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y}) \\ &= |(\mathbf{x}, Q^*Q\mathbf{y} - \mathbf{y})| \end{aligned}$$

and since $\mathbf{x}$ is arbitrary, it follows that for all $\mathbf{y}$,

$$Q^*Q\mathbf{y} - \mathbf{y} = \mathbf{0}$$

Thus

$$I = Q^*Q.$$

Similarly $QQ^* = I$. ■

12.4 The Tensor Product Of Two Vectors

Definition 12.4.1 *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 12.4.2 *Let X, Y, Z be inner product spaces. Then for α a scalar,*

$$(\alpha(y \otimes x))^* = \overline{\alpha} x \otimes y \tag{12.8}$$

$$(z \otimes y_1)(y_2 \otimes x) = (y_2, y_1) z \otimes x \tag{12.9}$$

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 12.8.

To verify 12.9, 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 12.9 give the same answer for every $u \in X$, it follows the two transformations are the same. ■

Definition 12.4.3 *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 12.4.4 *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 12.4.5 *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. ■

12.5 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 12.5.1 *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 12.2.4 on Page 383 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.

12.6 Fredholm Alternative Again

The best context in which to study the Fredholm alternative is in inner product spaces. This is done here.

Definition 12.6.1 *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 12.6.2 *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 12.2.4 on Page 383 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 12.6.3 *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 $A(V) = W$. ■

12.7 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 $\left\{1, x, x^2, x^3\right\}$ 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

$$\int_a^b q(x)\, y(x)\, z(x)\, dx = 0. \tag{12.10}$$

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, 12.10 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 $\mathbb{R}^4$ $\mathbf{x}, \mathbf{y}$ 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 $A^T A\mathbf{x} = A^T \mathbf{b}$.

12.8 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 12.8.1 *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*

$$P(\mathbf{u}_1, \cdots, \mathbf{u}_p) \equiv \left\{ \sum_{j=1}^{p} s_j \mathbf{u}_j : s_j \in [0,1] \right\}.$$

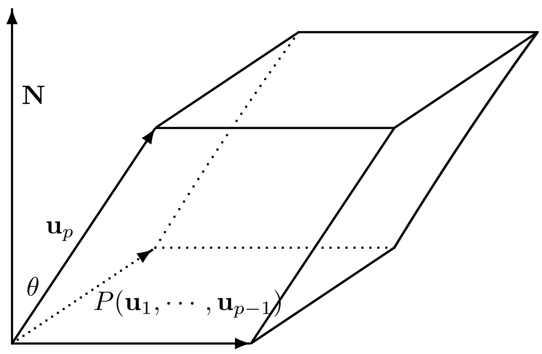
The volume of this parallelepiped is defined as

$$\text{volume of } P(\mathbf{u}_1, \cdots, \mathbf{u}_p) \equiv v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p)) \equiv (\det(\mathbf{u}_i \cdot \mathbf{u}_j))^{1/2}.$$

If the vectors 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 is a linear combination of the other rows and so from the properties of the determinant, the determinant of this matrix is indeed zero as it should be.

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 we want 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

$$v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p)) = |\mathbf{u}_p| \, |\cos(\theta)| \, v(P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})) \tag{12.11}$$

In the case where $p = 1$, the parallelepiped $P(\mathbf{v})$ consists of the single vector and the one dimensional volume should be $|\mathbf{v}| = \left(\mathbf{v}^T \mathbf{v}\right)^{1/2}$. Now having made this definition, I will show that this is the appropriate definition of p dimensional volume for every p.

Definition 12.8.2 *Let $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ be vectors. Then*

$$v(P(\mathbf{u}_1, \cdots, \mathbf{u}_p)) \equiv$$

$$\equiv \det \left(\begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_p^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p \end{pmatrix} \right)^{1/2}$$

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 12.11 holds.

Theorem 12.8.3 *With the above definition of volume, 12.11 holds.*

Proof: To check whether this is so, it is necessary to find $|\cos(\theta)|$. This involves finding the vector perpendicular to $P(\mathbf{u}_1, \cdots, \mathbf{u}_{p-1})$. Let $\{\mathbf{w}_1, \cdots, \mathbf{w}_p\}$ be 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 $k \leq p$. Such an orthonormal basis exists because of the Gram Schmidt procedure. First note that since $\{\mathbf{w}_k\}$ is an orthonormal basis for $\text{span}(\mathbf{u}_1, \cdots, \mathbf{u}_p)$,

$$\mathbf{u}_j = \sum_{k=1}^{p} (\mathbf{u}_j \cdot \mathbf{w}_k) \mathbf{w}_k$$

and if $i, j \leq k$

$$\mathbf{u}_j \cdot \mathbf{u}_i = \sum_{k=1}^{k} (\mathbf{u}_j \cdot \mathbf{w}_k)(\mathbf{u}_i \cdot \mathbf{w}_k)$$

Therefore, for each $k \leq p$

$$\det \left(\begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_k^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix} \right)$$

is the determinant of a matrix whose ij^{th} entry is

$$\mathbf{u}_i^T \mathbf{u}_j = \mathbf{u}_i \cdot \mathbf{u}_j = \sum_{r=1}^{k} (\mathbf{u}_i \cdot \mathbf{w}_r)(\mathbf{w}_r \cdot \mathbf{u}_j)$$

Thus this matrix is the product of the two $k \times k$ matrices, one which is the transpose of the other.

$$\begin{pmatrix} (\mathbf{u}_1 \cdot \mathbf{w}_1) & (\mathbf{u}_1 \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_1 \cdot \mathbf{w}_k) \\ (\mathbf{u}_2 \cdot \mathbf{w}_1) & (\mathbf{u}_2 \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_2 \cdot \mathbf{w}_k) \\ \vdots & \vdots & & \vdots \\ (\mathbf{u}_k \cdot \mathbf{w}_1) & (\mathbf{u}_k \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_k \cdot \mathbf{w}_k) \end{pmatrix} \cdot$$
$$\begin{pmatrix} (\mathbf{u}_1 \cdot \mathbf{w}_1) & (\mathbf{u}_2 \cdot \mathbf{w}_1) & \cdots & (\mathbf{u}_k \cdot \mathbf{w}_1) \\ (\mathbf{u}_1 \cdot \mathbf{w}_2) & (\mathbf{u}_2 \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_k \cdot \mathbf{w}_2) \\ \vdots & \vdots & & \vdots \\ (\mathbf{u}_1 \cdot \mathbf{w}_k) & (\mathbf{u}_2 \cdot \mathbf{w}_k) & \cdots & (\mathbf{u}_k \cdot \mathbf{w}_k) \end{pmatrix}$$

It follows

$$\det \left(\begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_k^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix} \right)$$

$$= \left(\det \begin{pmatrix} (\mathbf{u}_1 \cdot \mathbf{w}_1) & (\mathbf{u}_1 \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_1 \cdot \mathbf{w}_k) \\ (\mathbf{u}_2 \cdot \mathbf{w}_1) & (\mathbf{u}_2 \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_2 \cdot \mathbf{w}_k) \\ \vdots & \vdots & & \vdots \\ (\mathbf{u}_k \cdot \mathbf{w}_1) & (\mathbf{u}_k \cdot \mathbf{w}_2) & \cdots & (\mathbf{u}_k \cdot \mathbf{w}_k) \end{pmatrix} \right)^2$$

and so from the definition,

$$v \left(P \left(\mathbf{u}_1, \cdots, \mathbf{u}_k \right) \right) =$$

$$\left|\det\begin{pmatrix} (\mathbf{u}_1\cdot\mathbf{w}_1) & (\mathbf{u}_1\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_1\cdot\mathbf{w}_k) \\ (\mathbf{u}_2\cdot\mathbf{w}_1) & (\mathbf{u}_2\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_2\cdot\mathbf{w}_k) \\ \vdots & \vdots & & \vdots \\ (\mathbf{u}_k\cdot\mathbf{w}_1) & (\mathbf{u}_k\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_k\cdot\mathbf{w}_k) \end{pmatrix}\right|$$

Now consider the vector

$$\mathbf{N} \equiv \det\begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_p \\ (\mathbf{u}_1\cdot\mathbf{w}_1) & (\mathbf{u}_1\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_1\cdot\mathbf{w}_p) \\ \vdots & \vdots & & \vdots \\ (\mathbf{u}_{p-1}\cdot\mathbf{w}_1) & (\mathbf{u}_{p-1}\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_{p-1}\cdot\mathbf{w}_p) \end{pmatrix}$$

which results from formally expanding along the top row. Note that from what was just discussed,

$$v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1}\right)\right) = \pm A_{1p}$$

Now it follows from the formula for expansion of a determinant along the top row that for each $j \leq p-1$

$$\mathbf{N}\cdot\mathbf{u}_j = \sum_{k=1}^{p}\left(\mathbf{u}_j\cdot\mathbf{w}_k\right)\left(\mathbf{N}\cdot\mathbf{w}_k\right) = \sum_{k=1}^{p}\left(\mathbf{u}_j\cdot\mathbf{w}_k\right)A_{1k}$$

where A_{1k} is the $1k^{th}$ cofactor of the above matrix. Thus if $j \leq p-1$

$$\mathbf{N}\cdot\mathbf{u}_j = \det\begin{pmatrix} (\mathbf{u}_j\cdot\mathbf{w}_1) & (\mathbf{u}_j\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_j\cdot\mathbf{w}_p) \\ (\mathbf{u}_1\cdot\mathbf{w}_1) & (\mathbf{u}_1\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_1\cdot\mathbf{w}_p) \\ \vdots & \vdots & & \vdots \\ (\mathbf{u}_{p-1}\cdot\mathbf{w}_1) & (\mathbf{u}_{p-1}\cdot\mathbf{w}_2) & \cdots & (\mathbf{u}_{p-1}\cdot\mathbf{w}_p) \end{pmatrix} = 0$$

because the matrix has two equal rows while if $j = p$, the above discussion shows $\mathbf{N}\cdot\mathbf{u}_p$ equals $\pm v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_p\right)\right)$. Therefore, $\mathbf{N}$ points in the direction of the normal vector in the above picture or else it points in the opposite direction to this vector. From the geometric description of the dot product,

$$\left|\cos(\theta)\right| = \frac{\left|\mathbf{N}\cdot\mathbf{u}_p\right|}{\left|\mathbf{u}_p\right|\left|\mathbf{N}\right|}$$

and it follows

$$\left|\mathbf{u}_p\right|\left|\cos(\theta)\right| v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1}\right)\right) = \left|\mathbf{u}_p\right|\frac{\left|\mathbf{N}\cdot\mathbf{u}_p\right|}{\left|\mathbf{u}_p\right|\left|\mathbf{N}\right|} v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1}\right)\right)$$

$$= \frac{v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_p\right)\right)}{\left|\mathbf{N}\right|} v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1}\right)\right)$$

Now at this point, note that from the construction, $\mathbf{w}_p\cdot\mathbf{u}_k = 0$ whenever $k \leq p-1$ because $\mathbf{u}_k \in \operatorname{span}\left(\mathbf{w}_1,\cdots,\mathbf{w}_{p-1}\right)$. Therefore, $\left|\mathbf{N}\right| = \left|A_{1p}\right| = v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1}\right)\right)$ and so the above reduces to

$$\left|\mathbf{u}_p\right|\left|\cos(\theta)\right| v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1}\right)\right) = v\left(P\left(\mathbf{u}_1,\cdots,\mathbf{u}_p\right)\right). \blacksquare$$

The theorem shows that the only reasonable definition of p dimensional volume of a parallelepiped is the one given in the above definition.

12.9 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. Compare Definition 12.8.2 with the Binet Cauchy theorem, Theorem 3.3.14. What is the geometric meaning of the Binet Cauchy theorem in this context?

8. For W a subspace of V, W is said to have a complementary subspace [14] 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 10.3.4 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 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.

To see Chapter 13-15 download
Linear Algebra III Advanced topics