Linear Algebra
Hoffman & Kunze

2nd edition

Answers and Solutions to Problems and Exercises
Typos, comments and etc...

Gregory R. Grant
University of Pennsylvania
email: ggrant@upenn.edu

Julyl 2017

Note

This is one of the ultimate classic textbooks in mathematics. It will probably be read for generations to come. Yet I cannot find a comprehensive set of solutions online. Nor can I find lists of typos. Since I'm going through this book in some detail I figured I'd commit some of my thoughts and solutions to the public domain so others may have an easier time than I have finding supporting matericals for this book. Any book this classic should have such supporting materials readily avaiable.

If you find any mistakes in these notes, please do let me know at one of these email addresses:

ggrant543@gmail.com or greg@grant.org or ggrant@upenn.edu

The latex source file for this document is available at `http://greg.grant.org/hoffman_and_kunze.tex`. Use it wisely. And if you cannot manage to download the link because you included the period at the end of the sentence, then maybe you are not smart enough to be reading this book in the first place.

Chapter 1: Linear Equations

Section 1.1: Fields

Page 3. Hoffman and Kunze comment that the term "characteristic zero" is "strange." But the characteristic is the smallest n such that $n \cdot 1 = 0$. In a characteristic zero field the smallest such n is 0. This must be why they use the term "characteristic zero" and it doesn't seem that strange.

Section 1.2: Systems of Linear Equations

Page 5 Clarification: In Exercise 6 of this section they ask us to show, in the special case of two equations and two unknowns, that two homogeneous linear systems have the exact same solutions then they have the same row-reduced echelon form (we know the converse is always true by Theorem 3, page 7). Later in Exercise 10 of section 1.4 they ask us to prove it when there are two equations and three unknowns. But they never tell us whether this is true in general (for abitrary numbers of unknowns and equations). In fact is *is* true in general. This explanation was given on math.stackexchange:

> Solutions to the homogeneous system associated with a matrix is the same as determining the null space of the relevant matrix. The row space of a matrix is complementary to the null space. This is true not only for inner product spaces, and can be proved using the theory of non-degenerate symmetric bilinear forms.
>
> So if two matrices of the same order have exactly the same null space, they must also have exactly the same row space. In the row reduced echelon form the nonzero rows form a basis for the row space of the original matrix, and hence two matrices with the same row space will have the same row reduced echelon form.

Exercise 1: Verify that the set of complex numbers described in Example 4 is a subfield of $\mathbb{C}$.

Solution: Let $F = \{x + y\sqrt{2} \mid x, y \in \mathbb{Q}\}$. Then we must show six things:

1. 0 is in F
2. 1 is in F
3. If x and y are in F then so is $x + y$
4. If x is in F then so is $-x$
5. If x and y are in F then so is xy
6. If $x \neq 0$ is in F then so is x^{-1}

For 1, take $x = y = 0$. For 2, take $x = 1$, $y = 0$. For 3, suppose $x = a+b\sqrt{2}$ and $y = c+d\sqrt{2}$. Then $x+y = (a+c)+(b+d)\sqrt{2} \in F$. For 4, suppose $x = a + b\sqrt{2}$. Then $-x = (-a) + (-b)\sqrt{2} \in F$. For 5, suppose $x = a + b\sqrt{2}$ and $y = c + d\sqrt{2}$. Then

$xy = (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2} \in F$. For 6, suppose $x = a + b\sqrt{2}$ where at least one of a or b is not zero. Let $n = a^2 - 2b^2$. Let $y = a/n + (-b/n)\sqrt{2} \in F$. Then $xy = \frac{1}{n}(a + b\sqrt{2})(a - b\sqrt{2}) = \frac{1}{n}(a^2 - 2b^2) = 1$. Thus $y = x^{-1}$ and $y \in F$.

Exercise 2: Let F be the field of complex numbers. Are the following two systems of linear equations equivalent? If so, express each equation in each system as a linear combination of the equations in the other system.

$$\begin{aligned} x_1 - x_2 &= 0 & \qquad 3x_1 + x_2 &= 0 \\ 2x_1 + x_2 &= 0 & x_1 + x_2 &= 0 \end{aligned}$$

Solution: Yes the two systems are equivalent. We show this by writing each equation of the first system in terms of the second, and conversely.

$$\begin{aligned} 3x_1 + x_2 &= \frac{1}{3}(x_1 - x_2) + \frac{4}{3}(2x_1 + x_2) \\ x_1 + x_2 &= \frac{-1}{3}(x_1 - x_2) + \frac{2}{3}(2x_1 + x_2) \\ x_1 - x_2 &= (3x_1 + x_2) - 2(x_1 + x_2) \\ 2x_1 + x_2 &= \frac{1}{2}(3x_1 + x_2) + \frac{1}{2}(x_1 + x_2) \end{aligned}$$

Exercise 3: Test the following systems of equations as in Exercise 2.

$$\begin{aligned} -x_1 + x_2 + 4x_3 &= 0 & \qquad x_1 - x_3 &= 0 \\ x_1 + 3x_2 + 8x_3 &= 0 & x_2 + x_3 &= 0 \\ \tfrac{1}{2}x_1 + x_2 + \tfrac{5}{2}x_3 &= 0 \end{aligned}$$

Solution: Yes the two systems are equivalent. We show this by writing each equation of the first system in terms of the second, and conversely.

$$\begin{aligned} x_1 - x_3 &= \tfrac{-3}{4}(-x_1 + x_2 + 4x_3) + \tfrac{1}{4}(x_1 + 3x_3 + 8x_3) \\ x_2 + 3x_3 &= \tfrac{1}{4}(-x_1 + x_2 + 4x_3) + \tfrac{1}{4}(x_1 + 3x_3 + 8x_3) \end{aligned}$$

and

$$\begin{aligned} -x_1 + x_2 + 4x_3 &= -(x_1 - x_3) + (x_2 + 3x_3) \\ x_1 + 3x_2 + 8x_3 &= (x_1 - x_3) + 3(x_2 + 3x_3) \\ \tfrac{1}{2}x_1 + x_2 + \tfrac{5}{2}x_3 &= \tfrac{1}{2}(x_1 - x_3) + (x_2 + 3x_3) \end{aligned}$$

Exercise 4: Test the following systems as in Exercie 2.

$$\begin{aligned} 2x_1 + (-1 + i)x_2 \qquad\qquad + x_4 &= 0 & \qquad (1 + \tfrac{i}{2})x_1 + 8x_2 - ix_3 - x_4 &= 0 \\ 3x_2 - 2ix_3 + 5x_4 &= 0 & \tfrac{2}{3}x_1 - \tfrac{1}{2}x_2 + x_3 + 7x_4 &= 0 \end{aligned}$$

Solution: These systems are not equivalent. Call the two equations in the first system E_1 and E_2 and the equations in the second system E_1' and E_2'. Then if $E_2' = aE_1 + bE_2$ since E_2 does not have x_1 we must have $a = 1/3$. But then to get the coefficient of x_4 we'd need $7x_4 = \frac{1}{3}x_4 + 5bx_4$. That forces $b = \frac{4}{3}$. But if $a = \frac{1}{3}$ and $b = \frac{4}{3}$ then the coefficient of x_3 would have to be $-2i\frac{4}{3}$ which does not equal 1. Therefore the systems cannot be equivalent.

Exercise 5: Let F be a set which contains exactly two elements, 0 and 1. Define an addition and multiplication by the tables:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
0	0	1

Solution: We must check the nine conditions on pages 1-2:

1. An operation is commutative if the table is symmetric across the diagonal that goes from the top left to the bottom right. This is true for the addition table so addition is commutative.
2. There are eight cases. But if $x = y = z = 0$ or $x = y = z = 1$ then it is obvious. So there are six non-trivial cases. If there's exactly one 1 and two 0's then both sides equal 1. If there are exactly two 1's and one 0 then both sides equal 0. So addition is associative.
3. By inspection of the addition table, the element called 0 indeed acts like a zero, it has no effect when added to another element.
4. $1 + 1 = 0$ so the additive inverse of 1 is 1. And $0 + 0 = 0$ so the additive inverse of 0 is 0. In other words $-1 = 1$ and $-0 = 0$. So every element has an additive inverse.
5. As stated in 1, an operation is commutative if the table is symmetric across the diagonal that goes from the top left to the bottom right. This is true for the multiplication table so multiplication is commutative.
6. As with addition, there are eight cases. If $x = y = z = 1$ then it is obvious. Otherwise at least one of x, y or z must equal 0. In this case both $x(yz)$ and $(xy)z$ equal zero. Thus multiplication is associative.
7. By inspection of the multiplication table, the element called 1 indeed acts like a one, it has no effect when multiplied to another element.
8. There is only one non-zero element, 1. And $1 \cdot 1 = 1$. So 1 has a multiplicative inverse. In other words $1^{-1} = 1$.
9. There are eight cases. If $x = 0$ then clearly both sides equal zero. That takes care of four cases. If all three $x = y = z = 1$ then it is obvious. So we are down to three cases. If $x = 1$ and $y = z = 0$ then both sides are zero. So we're down to the two cases where $x = 1$ and one of y or z equals 1 and the other equals 0. In this case both sides equal 1. So $x(y + z) = (x + y)z$ in all eight cases.

Exercise 6: Prove that if two homogeneous systems of linear equations in two unknowns have the same solutions, then they are equivalent.

Solution: Write the two systems as follows:

$$\begin{array}{cc}
a_{11}x + a_{12}y = 0 & \qquad b_{11}x + b_{12}y = 0 \\
a_{21}x + a_{22}y = 0 & \qquad b_{21}x + b_{22}y = 0 \\
\vdots & \qquad \vdots \\
a_{m1}x + a_{m2}y = 0 & \qquad b_{m1}x + b_{m2}y = 0
\end{array}$$

Each system consists of a set of lines through the origin $(0, 0)$ in the x-y plane. Thus the two systems have the same solutions if and only if they either both have $(0, 0)$ as their only solution or if both have a single line $ux + vy - 0$ as their common solution. In the latter case all equations are simply multiples of the same line, so clearly the two systems are equivalent. So assume that both systems have $(0, 0)$ as their only solution. Assume without loss of generality that the first two equations in the first system give different lines. Then

$$\frac{a_{11}}{a_{12}} \neq \frac{a_{21}}{a_{22}} \tag{1}$$

We need to show that there's a (u, v) which solves the following system:

$$\begin{aligned}
a_{11}u + a_{12}v &= b_{i1} \\
a_{21}u + a_{22}v &= b_{i2}
\end{aligned}$$

Solving for u and v we get

$$u = \frac{a_{22}b_{i1} - a_{12}b_{i2}}{a_{11}a_{22} - a_{12}a_{21}}$$

$$v = \frac{a_{11}b_{i2} - a_{21}b_{i1}}{a_{11}a_{22} - a_{12}a_{12}}$$

By (1) $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Thus both u and v are well defined. So we can write any equation in the second system as a combination of equations in the first. Analogously we can write any equation in the first system in terms of the second.

Exercise 7: Prove that each subfield of the field of complex numbers contains every rational number.

Solution: Every subfield of $\mathbb{C}$ has characterisitc zero since if F is a subfield then $1 \in F$ and $n \cdot 1 = 0$ in F implies $n \cdot 1 = 0$ in $\mathbb{C}$. But we know $n \cdot 1 = 0$ in $\mathbb{C}$ implies $n = 0$. So $1, 2, 3, \dots$ are all distinct elements of F. And since F has additive inverses $-1, -2, -3, \dots$ are also in F. And since F is a field also $0 \in F$. Thus $\mathbb{Z} \subseteq F$. Now F has multiplicative inverses so $\pm\frac{1}{n} \in F$ for all natural numbers n. Now let $\frac{m}{n}$ be any element of $\mathbb{Q}$. Then we have shown that m and $\frac{1}{n}$ are in F. Thus their product $m \cdot \frac{1}{n}$ is in F. Thus $\frac{m}{n} \in F$. Thus we have shown all elements of $\mathbb{Q}$ are in F.

Exercise 8: Prove that each field of characteristic zero contains a copy of the rational number field.

Solution: Call the additive and multiplicative identities of F 0_F and 1_F respectively. Define n_F to be the sum of n 1_F's. So $n_F = 1_F + 1_F + \cdots + 1_F$ (n copies of 1_F). Define $-n_F$ to be the additive inverse of n_F. Since F has characteristic zero, if $n \neq m$ then $n_F \neq m_F$. For $m, n \in \mathbb{Z}$, $n \neq 0$, let $\frac{m}{n}_F = m_F \cdot n_F^{-1}$. Since F has characteristic zero, if $\frac{m}{n} \neq \frac{m'}{n'}$ then $\frac{m}{n}_F \neq \frac{m'}{n'}_F$. Therefore the map $\frac{m}{n} \mapsto \frac{m}{n}_F$ gives a one-to-one map from $\mathbb{Q}$ to F. Call this map h. Then $h(0) = 0_F$, $h(1) = 1_F$ and in general $h(x + y) = h(x) + h(y)$ and $h(xy) = h(x)h(y)$. Thus we have found a subset of F that is in one-to-one correspondence to $\mathbb{Q}$ and which has the same field structure as $\mathbb{Q}$.

Section 1.3: Matrices and Elementary Row Operations

Page 10, typo in proof of Theorem 4. Pargraph 2, line 6, it says $k_r \neq k$ but on the next line they call it k' instead of k_r. I think it's best to use k', because k_r is a more confusing notation.

Exercise 1: Find all solutions to the systems of equations

$$(1 - i)x_1 - ix_2 = 0$$
$$2x_1 + (1 - i)x_2 = 0.$$

Solution: The matrix of coefficients is

$$\begin{bmatrix} 1-i & -i \\ 2 & 1-i \end{bmatrix}.$$

Row reducing

$$\rightarrow \begin{bmatrix} 2 & 1-i \\ 1-i & -i \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1-i \\ 0 & 0 \end{bmatrix}$$

Thus $2x_1 + (1 - i)x_2 = 0$. Thus for any $x_2 \in \mathbb{C}$, $(\frac{1}{2}(i - 1)x_2, x_2)$ is a solution and these are all solutions.

Exercise 2: If

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{bmatrix}$$

find all solutions of $AX = 0$ by row-reducing A.

Solution:

$$\rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 2 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 7 & 1 \\ 0 & 8 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 1/7 \\ 0 & 8 & 2 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 3/7 \\ 0 & 1 & 1/7 \\ 0 & 0 & 6/7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 3/7 \\ 0 & 1 & 1/7 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 10 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus A is row-equivalent to the identity matrix. It follows that the only solution to the system is $(0, 0, 0)$.

Exercise 3: If
$$A = \begin{bmatrix} 6 & -4 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & 3 \end{bmatrix}$$
find all solutions of $AX = 2X$ and all solutions of $AX = 3X$. (The symbol cX denotes the matrix each entry of which is c times the corresponding entry of X.)

Solution: The system $AX = 2X$ is
$$\begin{bmatrix} 6 & -4 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
which is the same as
$$\begin{aligned} 6x - 4y &= 2x \\ 4x - 2y &= 2y \\ -x + 3z &= 2z \end{aligned}$$
which is equivalent to
$$\begin{aligned} 4x - 4y &= 0 \\ 4x - 4y &= 0 \\ -x + z &= 0 \end{aligned}$$
The matrix of coefficients is
$$\begin{bmatrix} 4 & -4 & 0 \\ 4 & -4 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$
which row-reduces to
$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$
Thus the solutions are all elements of F^3 of the form (x, x, x) where $x \in F$.

The system $AX = 3X$ is
$$\begin{bmatrix} 6 & -4 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 3\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
which is the same as
$$\begin{aligned} 6x - 4y &= 3x \\ 4x - 2y &= 3y \\ -x + 3z &= 3z \end{aligned}$$
which is equivalent to
$$\begin{aligned} 3x - 4y &= 0 \\ x - 2y &= 0 \\ -x &= 0 \end{aligned}$$

The matrix of coefficients is

$$\begin{bmatrix} 3 & -4 & 0 \\ 1 & -2 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

which row-reduces to

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus the solutions are all elements of F^3 of the form $(0, 0, z)$ where $z \in F$.

Exercise 4: Find a row-reduced matrix which is row-equivalent to

$$A = \begin{bmatrix} i & -(1+i) & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{bmatrix}.$$

Solution:

$$A \to \begin{bmatrix} 1 & -2 & 1 \\ i & -(1+i) & 0 \\ 1 & 2i & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1+i & -i \\ 0 & 2+2i & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & \frac{1-i}{2} \\ 0 & 2+2i & -2 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & \frac{i-1}{2} \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & i \\ 0 & 1 & \frac{i-1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Exercise 5: Prove that the following two matrices are *not* row-equivalent:

$$\begin{bmatrix} 2 & 0 & 0 \\ a & -1 & 0 \\ b & c & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 2 \\ -2 & 0 & -1 \\ 1 & 3 & 5 \end{bmatrix}.$$

Solution: Call the first matrix A and the second matrix B. The matrix A is row-equivalent to

$$A' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the matrix B is row-equivalent to

$$B' = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 3/2 \\ 0 & 0 & 0 \end{bmatrix}.$$

By Theorem 3 page 7 $AX = 0$ and $A'X = 0$ have the same solutions. Similarly $BX = 0$ and $B'X = 0$ have the same solutions. Now if A and B are row-equivalent then A' and B' are row equivalent. Thus if A and B are row equivalent then $A'X = 0$ and $B'X = 0$ must have the same solutions. But $B'X = 0$ has infinitely many solutions and $A'X = 0$ has only the trivial solution $(0, 0, 0)$. Thus A and B cannot be row-equivalent.

Exercise 6: Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be a 2×2 matrix with complex entries. Suppose that A is row-reduced and also that $a + b + c + d = 0$. Prove that there are exactly three such matrices.

Solution: Case $a \neq 0$: Then to be in row-reduced form it must be that $a = 1$ and $A = \begin{bmatrix} 1 & b \\ c & d \end{bmatrix}$ which implies $c = 0$, so $A = \begin{bmatrix} 1 & b \\ 0 & d \end{bmatrix}$. Suppose $d \neq 0$. Then to be in row-reduced form it must be that $d = 1$ and $b = 0$, so $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ which implies $a + b + c + d \neq 0$. So it must be that $d = 0$, and then it follows that $b = -1$. So $a \neq 0 \Rightarrow A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$.

Case $a = 0$: Then $A = \begin{bmatrix} 0 & b \\ c & d \end{bmatrix}$. If $b \neq 0$ then b must equal 1 and $A = \begin{bmatrix} 0 & 1 \\ c & d \end{bmatrix}$ which forces $d = 0$. So $A = \begin{bmatrix} 0 & 1 \\ c & 0 \end{bmatrix}$ which implies (since $a + b + c + d = 0$) that $c = -1$. But c cannot be -1 in row-reduced form. So it must be that $b = 0$. So $A = \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}$. If $c \neq 0$ then $c = 1$, $d = -1$ and $A = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$. Otherwise $c = 0$ and $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Thus the three possibilities are:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}.$$

Exercise 7: Prove that the interchange of two rows of a matrix can be accomplished by a finite sequence of elementary row operations of the other two types.

Solution: Write the matrix as

$$\begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_n \end{bmatrix}.$$

WOLOG we'll show how to exchange rows R_1 and R_2. First add R_2 to R_1:

$$\begin{bmatrix} R_1 + R_2 \\ R_2 \\ R_3 \\ \vdots \\ R_n \end{bmatrix}.$$

Next subtract row one from row two:

$$\begin{bmatrix} R_1 + R_2 \\ -R_1 \\ R_3 \\ \vdots \\ R_n \end{bmatrix}.$$

Next add row two to row one again

$$\begin{bmatrix} R_2 \\ -R_1 \\ R_3 \\ \vdots \\ R_n \end{bmatrix}.$$

Finally multiply row two by -1:

$$\begin{bmatrix} R_2 \\ R_1 \\ R_3 \\ \vdots \\ R_n \end{bmatrix}.$$

Exercise 8: Consider the system of equations $AX = 0$ where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is a 2×2 matrix over the field F. Prove the following:

(a) If every entry of A is 0, then every pair (x_1, x_2) is a solution of $AX = 0$.

(b) If $ad - bc \neq 0$, the system $AX = 0$ has only the trivial solution $x_1 = x_2 = 0$.

(c) If $ad - bc = 0$ and some entry of A is different from 0, then there is a solution (x_1^0, x_2^0) such that (x_1, x_2) is a solution if and only if there is a scalar y such that $x_1 = yx_1^0$, $x_2 = yx_2^0$.

Solution:

(a) In this case the system of equations is

$$\begin{aligned} 0 \cdot x_1 + 0 \cdot x_2 &= 0 \\ 0 \cdot x_1 + 0 \cdot x_2 &= 0 \end{aligned}$$

Clearly any (x_1, x_2) satisfies this system since $0 \cdot x = 0 \; \forall \; x \in F$.

(b) Let $(u, v) \in F^2$. Consider the system:

$$\begin{aligned} a \cdot x_1 + b \cdot x_2 &= u \\ c \cdot x_1 + d \cdot x_2 &= v \end{aligned}$$

If $ad - bc \neq 0$ then we can solve for x_1 and x_2 explicitly as

$$x_1 = \frac{du - bv}{ad - bc} \qquad x_2 = \frac{av - cu}{ad - bc}.$$

Thus there's a unique solution for all (u, v) and in partucular when $(u, v) = (0, 0)$.

(c) Assume WOLOG that $a \neq 0$. Then $ad - bc = 0 \Rightarrow d = \frac{cb}{a}$. Thus if we multiply the first equation by $\frac{c}{a}$ we get the second equation. Thus the two equations are redundant and we can just consider the first one $ax_1 + bx_2 = 0$. Then any solution is of the form $(-\frac{b}{a}y, y)$ for arbitrary $y \in F$. Thus letting $y = 1$ we get the solution $(-b/a, 1)$ and the arbitrary solution is of the form $y(-b/a, 1)$ as desired.

Section 1.4: Row-Reduced Echelon Matrices

Exercise 1: Find all solutions to the following system of equations by row-reducing the coefficient matrix:

$$\begin{aligned} \frac{1}{3}x_1 + 2x_2 - 6x_3 &= 0 \\ -4x_1 \qquad\quad + 5x_3 &= 0 \\ -3x_1 + 6x_2 - 13x_3 &= 0 \\ -\frac{7}{3}x_1 + 2x_2 - \frac{8}{3}x_3 &= 0 \end{aligned}$$

Solution: The coefficient matrix is

$$\begin{bmatrix} \frac{1}{3} & 2 & -6 \\ -4 & 0 & 5 \\ -3 & 6 & -13 \\ -\frac{7}{3} & 2 & -\frac{8}{3} \end{bmatrix}$$

This reduces as follows:

$$\rightarrow \begin{bmatrix} 1 & 6 & -18 \\ -4 & 0 & 5 \\ -3 & 6 & -13 \\ -7 & 6 & -8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 6 & -18 \\ 0 & 24 & -67 \\ 0 & 24 & -67 \\ 0 & 48 & -134 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 6 & -18 \\ 0 & 24 & -67 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 6 & -18 \\ 0 & 1 & -67/24 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -5/4 \\ 0 & 1 & -67/24 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus

$$x - \frac{5}{4}z = 0$$

$$y - \frac{67}{24}z = 0$$

Thus the general solution is $(\frac{5}{4}z, \frac{67}{24}z, z)$ for arbitrary $z \in F$.

Exercise 2: Find a row-reduced echelon matrix which is row-equivalent to

$$A = \begin{bmatrix} 1 & -i \\ 2 & 2 \\ u & 1+i \end{bmatrix}.$$

What are the solutions of $AX = 0$?

Solution: A row-reduces as follows:

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

Thus the only solution to $AX = 0$ is $(0, 0)$.

Exercise 3: Describe explicitly all 2×2 row-reduced echelon matrices.

Solution:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & x \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Exercise 4: Consider the system of equations

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + 2x_3 &= 1 \\ x_1 - 3x_2 + 4x_3 &= 2 \end{aligned}$$

Does this system have a solution? If so, describe explicitly all solutions.

Solution: The augmented coefficient matrix is

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 2 & 0 & 2 & 1 \\ 1 & -3 & 4 & 2 \end{array}\right]$$

We row reduce it as follows:

$$\rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & 2 & -2 & -1 \\ 0 & -2 & 2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & -1/2 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1/2 \\ 0 & 1 & -1 & -1/2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus the system is equivalent to

$$\begin{aligned} x_1 + x_3 &= 1/2 \\ x_2 - x_3 &= -1/2 \end{aligned}$$

Thus the solutions are parameterized by x_3. Setting $x_3 = c$ gives $x_1 = 1/2 - c$, $x_2 = c - 1/2$. Thus the general solution is

$$\left(\tfrac{1}{2} - c,\ \ c - \tfrac{1}{2},\ \ c\right)$$

for $c \in \mathbb{R}$.

Exercise 5: Give an example of a system of two linear equations in two unkowns which has no solutions.

Solution:

$$x + y = 0$$

$$x + y = 1$$

Exercise 6: Show that the system

$$\begin{aligned} x_1 - 2x_2 + x_3 + 2x_4 &= 1 \\ x_1 + x_2 - x_3 + x_4 &= 2 \\ x_1 + 7x_2 - 5x_3 - x_4 &= 3 \end{aligned}$$

has no solution.

Solution: The augmented coefficient matrix is as follows

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 1 \\ 1 & 1 & -1 & 1 & 2 \\ 1 & 7 & -5 & -1 & 3 \end{array}\right]$$

This row reduces as follows:

$$\rightarrow \left[\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 1 \\ 0 & 3 & -2 & -1 & 1 \\ 0 & 9 & -6 & -3 & 2 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 1 \\ 0 & 3 & -2 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

At this point there's no need to continue because the last row says $0x_1 + 0x_2 + 0x_3 + 0x_4 = -1$. But the left hand side of this equation is zero so this is impossible.

Exercise 7: Find all solutions of

$$\begin{aligned} 2x_1 - 3x_2 - 7x_3 + 5x_4 + 2x_5 &= -2 \\ x_1 - 2x_2 - 4x_3 + 3x_4 + x_5 &= -2 \\ 2x_1 \qquad - 4x_3 + 2x_4 + x_5 &= 3 \\ x_1 - 5x_2 - 7x_3 + 6x_4 + 2x_5 &= -7 \end{aligned}$$

Solution: The augmented coefficient matrix is

$$\left[\begin{array}{ccccc|c} 2 & -3 & -7 & 5 & 2 & -2 \\ 1 & -2 & -4 & 3 & 1 & -2 \\ 2 & 0 & -4 & 2 & 1 & 3 \\ 1 & -5 & -7 & 6 & 2 & -7 \end{array}\right]$$

We row-reduce it as follows

$$\rightarrow \left[\begin{array}{ccccc|c} 1 & -2 & -4 & 3 & 1 & -2 \\ 2 & -3 & -7 & 5 & 2 & -2 \\ 2 & 0 & -4 & 2 & 1 & 3 \\ 1 & -5 & -7 & 6 & 2 & -7 \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & -2 & -4 & 3 & 1 & -2 \\ 0 & 1 & 1 & -1 & 0 & 2 \\ 0 & 4 & 4 & -4 & -1 & 7 \\ 0 & -3 & -3 & 3 & 1 & -5 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc|c} 1 & 0 & -2 & 1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & 0 & -2 & 1 & 0 & 1 \\ 0 & 1 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Thus

$$\begin{aligned} x_1 - 2x_3 + x_4 &= 1 \\ x_2 + x_3 - x_4 &= 2 \\ x_5 &= 1 \end{aligned}$$

Thus the general solution is given by $(1 + 2x_3 - x_4, 2 + x_4 - x_3, x_3, x_4, 1)$ for arbitrary $x_3, x_4 \in F$.

Exercise 8: Let

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{bmatrix}.$$

For which (y_1, y_2, y_3) does the system $AX = Y$ have a solution?

Solution: The matrix A is row-reduced as follows:

$$\rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 7 & 1 \\ 0 & 8 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 7 & 1 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 1 \\ 0 & 7 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus for every (y_1, y_2, y_3) there is a (unique) solution.

Exercise 9: Let

$$\begin{bmatrix} 3 & -6 & 2 & -1 \\ -2 & 4 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix}.$$

For which (y_1, y_2, y_3, y_4) does the system of equations $AX = Y$ have a solution?

Solution: We row reduce as follows

$$\left[\begin{array}{cccc|c} 3 & -6 & 2 & -1 & y_1 \\ -2 & 4 & 1 & 3 & y_2 \\ 0 & 0 & 1 & 1 & y_3 \\ 1 & -2 & 1 & 0 & y_4 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & y_4 \\ 3 & -6 & 2 & -1 & y_1 \\ -2 & 4 & 1 & 3 & y_2 \\ 0 & 0 & 1 & 1 & y_3 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & y_4 \\ 0 & 0 & -1 & -1 & y_1 - 3y_4 \\ 0 & 0 & 3 & 3 & y_2 + 2y_4 \\ 0 & 0 & 1 & 1 & y_3 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & y_4 \\ 0 & 0 & 0 & 0 & y_1 - 3y_4 + y_3 \\ 0 & 0 & 0 & 0 & y_2 + 2y_4 + 3y_3 \\ 0 & 0 & 1 & 1 & y_3 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & y_4 \\ 0 & 0 & 1 & 1 & y_3 \\ 0 & 0 & 0 & 0 & y_1 - 3y_4 + y_3 \\ 0 & 0 & 0 & 0 & y_2 + 2y_4 + 3y_3 \end{array}\right]$$

Thus (y_1, y_2, y_3, y_4) must satisfy

$$\begin{aligned} y_1 + y_3 - 3y_4 &= 0 \\ y_2 + 3y_3 + 2y_4 &= 0 \end{aligned}$$

The matrix for this system is

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

of which the general solution is $(-y_3 + 3y_4, -3y_3 - 2y_4, y_3, y_4)$ for arbitrary $y_3, y_4 \in F$. These are the only (y_1, y_2, y_3, y_4) for which the system $AX = Y$ has a solution.

Exercise 10: Suppose R and R' are 2×3 row-reduced echelon matrices and that the system $RX = 0$ and $R'X = 0$ have exactly the same solutions. Prove that $R = R'$.

Solution: There are seven possible 2×3 row-reduced echelon matrices:

$$R_1 = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \tag{2}$$

$$R_2 = \left[\begin{array}{ccc} 1 & 0 & a \\ 0 & 1 & b \end{array}\right] \tag{3}$$

$$R_3 = \left[\begin{array}{ccc} 1 & a & 0 \\ 0 & 0 & 1 \end{array}\right] \tag{4}$$

$$R_4 = \left[\begin{array}{ccc} 1 & a & b \\ 0 & 0 & 0 \end{array}\right] \tag{5}$$

$$R_5 = \left[\begin{array}{ccc} 0 & 1 & a \\ 0 & 0 & 0 \end{array}\right] \tag{6}$$

$$R_6 = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \tag{7}$$

$$R_7 = \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right] \tag{8}$$

We must show that no two of these have exactly the same solutions. For the first one R_1, any (x, y, z) is a solution and that's not the case for any of the other R_i's. Consider next R_7. In this case $z = 0$ and x and y can be anything. We can have $z \neq 0$ for R_2, R_3 and R_5. So the only ones R_7 could share solutions with are R_3 or R_6. But both of those have restrictions on x and/or y so the solutions cannot be the same. Also R_3 and R_6 cannot have the same solutions since R_6 forces $y = 0$ while R_3 does not.

Thus we have shown that if two R_i's share the same solutions then they must be among R_2, R_4, and R_5.

The solutions for R_2 are $(-az, -bz, z)$, for z arbitrary. The solutions for R_4 are $(-a'y-b'z, y, z)$ for y, z arbitrary. Thus $(-b', 0, 1)$ is a solution for R_4. Suppose this is also a solution for R_2. Then $z = 1$ so it is of the form $(-a, -b, 1)$ and it must be that $(-b', 0, 1) = (-a, -b, 1)$. Comparing the second component implies $b = 0$. But if $b = 0$ then R_2 implies $y = 0$. But R_4 allows for arbitrary y. Thus R_2 and R_4 cannot share the same solutions.

The solutions for R_2 are $(-az, -bz, z)$, for z arbitrary. The solutions for R_5 are $(x, -a'z, z)$ for x, z arbitrary. Thus $(0, -a', 1)$ is a solution for R_5. As before if this is a solution of R_2 then $a = 0$. But if $a = 0$ then R_2 forces $x = 0$ while in R_5 x can be arbitrary. Thus R_2 and R_5 cannot share the same solutions.

The solutions for R_4 are $(-ay - bz, y, z)$ for y, z arbitrary. The solutions for R_5 are $(x, -a'z, z)$ for x, z arbitrary. Thus setting $x = 1$, $z = 0$ gives $(1, 0, 0)$ is a solution for R_5. But this cannot be a solution for R_4 since if $y = z = 0$ then first component must also be zero.

Thus we have shown that no two R_i and R_j have the same solutions unless $i = j$.

NOTE: This fact is actually true in general not just for 2×3 (search for "1832109" on math.stackexchange).

Section 1.5: Matrix Multiplication

Page 18: Typo in the last paragraph where it says "the columns of B are the $1 \times n$ matrices ..." it should be $n \times 1$ not $1 \times n$.

Page 20: Typo in the Definition, it should say "An $m \times m$ matrix is said to be an **elementary matrix**"... Otherwise it doesn't make sense that you can obtain an $m \times n$ matrix from an $m \times m$ matrix by an elementary row operation unless $m = n$.

Exercise 1: Let

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \quad C = [1 \quad -1].$$

Compute ABC and CAB.

Solution:

$$AB = \begin{bmatrix} 4 \\ 4 \end{bmatrix},$$

so

$$ABC = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \cdot [1 \quad -1] = \begin{bmatrix} 4 & -4 \\ 4 & -4 \end{bmatrix}.$$

and

$$CBA = [1 \quad -1] \cdot \begin{bmatrix} 4 \\ 4 \end{bmatrix} = [0].$$

Exercise 2: Let

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -2 \\ 1 & 3 \\ 4 & 4 \end{bmatrix}.$$

Verify directly that $A(AB) = A^2B$.

Solution:

$$A^2 = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & -1 & 1 \\ 5 & -2 & 3 \\ 6 & -3 & 4 \end{bmatrix}.$$

And

$$AB = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & -2 \\ 1 & 3 \\ 4 & 4 \end{bmatrix}$$
$$= \begin{bmatrix} 5 & -1 \\ 8 & 0 \\ 10 & -2 \end{bmatrix}.$$

Thus

$$A^2B = \begin{bmatrix} 2 & -1 & 1 \\ 5 & -2 & 3 \\ 6 & -3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & -2 \\ 1 & 3 \\ 4 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & -3 \\ 20 & -4 \\ 25 & -5 \end{bmatrix}. \tag{9}$$

And

$$A(AB) = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 & -1 \\ 8 & 0 \\ 10 & -2 \end{bmatrix}$$

$$\begin{bmatrix} 7 & -3 \\ 20 & -4 \\ 25 & -5 \end{bmatrix}. \tag{10}$$

Comparing (9) and (10) we see both calculations result in the same matrix.

Exercise 3: Find two different 2×2 matrices A such that $A^2 = 0$ but $A \neq 0$.

Solution:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

are two such matrices.

Exercise 4: For the matrix A of Exercise 2, find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2 E_1 A = I.$$

Solution:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix}$$

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 3 & 0 & 1 \end{bmatrix}$$

$$E_2(E_1A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 3 & 0 \end{bmatrix}$$

$$E_3(E_2E_1A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1/2 \\ 0 & 3 & 0 \end{bmatrix}$$

$$E_4(E_3E_2E_1A) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1/2 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 3 & 0 \end{bmatrix}$$

$$E_5(E_4E_3E_2E_1A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 3/2 \end{bmatrix}$$

$$E_6(E_5E_4E_3E_2E_1A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2/3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 3/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_7(E_6E_5E_4E_3E_2E_1A) = \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_8(E_7E_6E_5E_4E_3E_2E_1A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exercise 5: Let

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ -4 & 4 \end{bmatrix}.$$

Is there a matrix C such that $CA = B$?

Solution: To find such a $C = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$ we must solve the equation

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -4 & 4 \end{bmatrix}.$$

This gives a system of equations

$$\begin{aligned} a + 2b + c &= 3 \\ -a + 2b &= 1 \\ d + 2e + f &= -4 \\ -d + 2e &= 4 \end{aligned}$$

We row-reduce the augmented coefficient matrix

$$\left[\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 0 & 0 & 3 \\ -1 & 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 2 & 0 & 4 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & 1/2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1/4 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1/2 & -4 \\ 0 & 0 & 0 & 0 & 1 & 1/4 & 0 \end{array}\right].$$

Setting $c = f = 4$ gives the solution

$$C = \begin{bmatrix} -1 & 0 & 4 \\ -6 & -1 & 4 \end{bmatrix}.$$

Checking:

$$\begin{bmatrix} -1 & 0 & 4 \\ -6 & -1 & 4 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -4 & 4 \end{bmatrix}.$$

Exercise 6: Let A be an $m \times n$ matrix and B an $n \times k$ matrix. Show that the columns of $C = AB$ are linear combinations of the columns of A. If $\alpha_1, \ldots, \alpha_n$ are the columns of A and $\gamma_1, \ldots, \gamma_k$ are the columns of C then

$$\gamma_j = \sum_{r=1}^{n} B_{rj}\alpha_r.$$

Solution: The ij-th entry of AB is $\sum_{r=1}^{k} A_{ir}B_{rj}$. Since the term B_{rj} is independent of i, we can view the sum independent of i as $\sum_{r=1}^{n} B_{rj}\alpha_r$ where α_r is the r-th column of A. I'm not sure what more to say, this is pretty immediately obvious from the definition of matrix multiplication.

Exercise 7: Let A and B be 2×2 matrices such that $AB = I$. Prove that $BA = I$.

Solution: Suppose

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}.$$

$$AB = \begin{bmatrix} ax+bz & ay+bw \\ cx+dz & cy+dw \end{bmatrix}.$$

Then $AB = I$ implies the following system in u, r, s, t has a solution

$$\begin{aligned} au + bs &= 1 \\ cu + ds &= 0 \\ ar + bt &= 0 \\ cr + dt &= 1 \end{aligned}$$

because (x, y, z, w) is one such solution. The augmented coefficient matrix of this system is

$$\left[\begin{array}{cccc|c} a & 0 & b & 0 & 1 \\ c & 0 & d & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & c & 0 & d & 1 \end{array}\right]. \tag{11}$$

As long as $ad - bc \neq 0$ this system row-reduces to the following row-reduced echelon form

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & d/(ad-bc) \\ 0 & 1 & 0 & 0 & -b/(ad-bc) \\ 0 & 0 & 1 & 0 & -c/(ad-bc) \\ 0 & 0 & 0 & 1 & a/(ad-bc) \end{array}\right]$$

Thus we see that necessarily $x = d/(ad-bc)$, $y = -b/(ad-bc)$, $z = -c/(ad-bc)$, $w = a/(ad-bc)$. Thus

$$B = \begin{bmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{bmatrix}.$$

Now it's a simple matter to check that

$$\begin{bmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The loose end is that we assumed $ad - bc \neq 0$. To tie up this loose end we must show that if $AB = I$ then necessarily $ad - bc \neq 0$. Suppose that $ad - bc = 0$. We will show there is no solution to (11), which contradicts the fact that (x, y, z, w) is a solution. If $a = b = c = d = 0$ then obviously $AB \neq I$. So suppose WOLOG that $a \neq 0$ (because by elementary row operations we can move any of the four elements to be the top left entry). Subtracting $\frac{c}{a}$ times the 3rd row from the 4th row of (11) gives

$$\left[\begin{array}{cccc|c} a & 0 & b & 0 & 1 \\ c & 0 & d & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & c - \frac{c}{a}a & 0 & d - \frac{c}{a}b & 1 \end{array}\right].$$

Now $c - \frac{c}{a}a = 0$ and since $ad - bc = 0$ also $d - \frac{c}{a}b = 0$. Thus we get

$$\left[\begin{array}{cccc|c} a & 0 & b & 0 & 1 \\ c & 0 & d & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

and it follows that (11) has no solution.

Exercise 8: Let

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

be a 2×2 matrix. We inquire when it is possible to find 2×2 matrices A and B such that $C = AB - BA$. Prove that such matrices can be found if and only fi $C_{11} + C_{22} = 0$.

Solution: We want to know when we can solve for a, b, c, d, x, y, z, w such that

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

The right hand side is equal to

$$\begin{bmatrix} bz - cy & ay + bw - bx - dy \\ cx + dz - az - cw & cy - bz \end{bmatrix}$$

Thus the question is equivalent to asking: when can we choose a, b, c, d so that the following system has a solution for x, y, z, w

$$\begin{aligned} bz - cy &= c_{11} \\ ay + bw - bx - dy &= c_{12} \\ cx + dz - az - cw &= c_{21} \\ cy - bz &= c_{22} \end{aligned} \tag{12}$$

The augmented coefficient matrix for this system is

$$\left[\begin{array}{cccc|c} 0 & -c & b & 0 & c_{11} \\ -b & a-d & 0 & b & c_{12} \\ c & 0 & d-a & -c & c_{21} \\ 0 & c & -b & 0 & c_{22} \end{array}\right]$$

This matrix is row-equivalent to

$$\left[\begin{array}{cccc|c} 0 & -c & b & 0 & c_{11} \\ -b & a-d & 0 & b & c_{12} \\ c & 0 & d-a & -c & c_{21} \\ 0 & 0 & 0 & 0 & c_{11} + c_{22} \end{array}\right]$$

from which we see that necessarily $c_{11} + c_{22} = 0$.

Suppose conversely that $c_{11} + c_{22} = 0$. We want to show $\exists\, A, B$ such that $C = AB - BA$.

We first handle the case when $c_{11} = 0$. We know $c_{11} + c_{22} = 0$ so also $c_{22} = 0$. So C is in the form

$$\begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \end{bmatrix}.$$

In this case let

$$A = \begin{bmatrix} 0 & c_{12} \\ -c_{21} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then

$$
\begin{aligned}
& AB - BA \\
&= \begin{bmatrix} 0 & c_{12} \\ -c_{21} & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & c_{12} \\ -c_{21} & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ c_{21} & 0 \end{bmatrix} - \begin{bmatrix} 0 & -c_{12} \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \end{bmatrix} \\
&= C.
\end{aligned}
$$

So we can assume going forward that $c_{11} \neq 0$. We want to show the system (12) can be solved. In other words we have to find a, b, c, d such that the system has a solution in x, y, z, w. If we assume $b \neq 0$ and $c \neq 0$ then this matrix row-reduces to the following row-reduced echelon form

$$
\left[\begin{array}{cccc|c}
1 & 0 & (d-a)/c & -1 & \frac{d-a}{bc}c_{11} - \frac{c_{12}}{b} \\
0 & 1 & -d/c & 0 & -c_{11}/c \\
0 & 0 & 0 & 0 & -\frac{c_{11}}{b}(d-a) + c_{21} + \frac{c_{12}c}{b} \\
0 & 0 & 0 & 0 & c_{11} + c_{22}
\end{array}\right]
$$

We see that necessarily

$$-\frac{c_{11}}{b}(d-a) + c_{21} + \frac{c_{12}c}{b} = 0.$$

Since $c_{11} \neq 0$, we can set $a = 0$, $b = c = 1$ and $d = \frac{c_{12}+c_{21}}{c_{11}}$. Then the system can be solved and we get a solution for any choice of z and w. Setting $z = w = 0$ we get $x = c_{21}$ and $y = c_{11}$.

Summarizing, if $c_{11} \neq 0$ then:

$$a = 0$$

$$b = 1$$

$$c = 1$$

$$d = (c_{12} + c_{21})/c_{11}$$

$$x = c_{21}$$

$$y = c_{11}$$

$$z = 0$$

$$w = 0$$

For example if $C = \begin{bmatrix} 2 & 1 \\ 3 & -2 \end{bmatrix}$ then $A = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -2 \\ 0 & 0 \end{bmatrix}$. Checking:

$$\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 3 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & -2 \end{bmatrix}.$$

Section 1.6: Invertible Matrices

Exercise 1: Let

$$A=\begin{bmatrix}1&2&1&0\\-1&0&3&5\\1&-2&1&1\end{bmatrix}.$$

Find a row-reduced echelon matrix R which is row-equivalent to A and an invertible 3×3 matrix P such that $R=PA$.

Solution: As in Exercise 4, Section 1.5, we row reduce and keep track of the elementary matrices involved. It takes nine steps to put A in row-reduced form resulting in the matrix

$$P=\begin{bmatrix}3/8&-1/4&3/8\\1/4&0&-1/4\\1/8&1/4&1/8\end{bmatrix}.$$

Exercise 2: Do Exercise 1, but with

$$A=\begin{bmatrix}2&0&i\\1&-3&-i\\i&1&1\end{bmatrix}.$$

Solution: Same story as Exercise 1, we get to the identity matrix in nine elementary steps. Multiplying those nine elementary matrices together gives

$$P=\begin{bmatrix}1/3&-\frac{29+3i}{30}&\frac{1-3i}{10}\\0&-\frac{3+i}{10}&\frac{1-3i}{10}\\-i/3&\frac{3+i}{15}&\frac{3+i}{5}\end{bmatrix}.$$

Exercise 3: For each of the two matrices

$$\begin{bmatrix}2&5&-1\\4&-1&2\\6&4&1\end{bmatrix},\quad\begin{bmatrix}1&-1&2\\3&2&4\\0&1&-2\end{bmatrix}$$

use elementary row operations to discover whether it is invertible, and to find the inverse in case it is.

Solution: For the first matrix we row-reduce the augmented matrix as follows:

$$\left[\begin{array}{ccc|ccc}2&5&-1&1&0&0\\4&-1&2&0&1&0\\6&4&1&0&0&1\end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc}2&5&-1&1&0&0\\0&-11&4&-2&1&0\\0&-11&4&-3&0&1\end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc}2&5&-1&1&0&0\\0&-11&4&-2&1&0\\0&0&0&-1&-1&1\end{array}\right]$$

At this point we see that the matrix is not invertible since we have obtained an entire row of zeros.

For the second matrix we row-reduce the augmented matrix as follows:

$$\left[\begin{array}{ccc|ccc}1&-1&2&1&0&0\\3&2&4&0&1&0\\0&1&-2&0&0&1\end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 5 & -2 & -3 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 5 & -2 & -3 & 1 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 8 & -3 & 1 & -5 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -3/8 & 1/8 & -5/8 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -3/4 & 1/4 & -1/4 \\ 0 & 0 & 1 & -3/8 & 1/8 & -5/8 \end{array}\right]$$

Thus the inverse matrix is

$$\left[\begin{array}{ccc} 1 & 0 & 1 \\ -3/4 & 1/4 & -1/4 \\ 3/8 & 1/8 & -5/8 \end{array}\right]$$

Exercise 4: Let

$$A = \left[\begin{array}{ccc} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{array}\right].$$

For which X does there exist a scalar c such that $AX = cX$?

Solution:

$$\left[\begin{array}{ccc} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{array}\right]\left[\begin{array}{c} x \\ y \\ z \end{array}\right] = c\left[\begin{array}{c} x \\ y \\ z \end{array}\right]$$

implies

$$5x = cx \tag{13}$$
$$x + 5y = cy \tag{14}$$
$$y + 5z = cz \tag{15}$$

Now if $c \neq 5$ then (13) implies $x = 0$, and then (14) implies $y = 0$, and then (15) implies $z = 0$. So it is true for $\left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right]$ with $c = 0$.

If $c = 5$ then (14) implies $x = 0$ and (15) implies $y = 0$. So if $c = 5$ any such vector must be of the form $\left[\begin{array}{c} 0 \\ 0 \\ z \end{array}\right]$ and indeed any such vector works with $c = 5$.

So the final answer is any vector of the form $\left[\begin{array}{c} 0 \\ 0 \\ z \end{array}\right]$.

Exercise 5: Discover whether

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

is invertible, and find A^{-1} if it exists.

Solution: We row-reduce the augmented matrix as follows:

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 2 & 3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 3 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/3 & -1/3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1/4 \end{array}\right]$$

Thus the A does have an inverse and

$$A^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1/3 & -1/3 \\ 0 & 0 & 0 & 1/4 \end{bmatrix}.$$

Exercise 6: Suppose A is a 2×1 matrix and that B is a 1×2 matrix. Prove that $C = AB$ is not invertible.

Solution: Write $A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $B = [b_1 \ \ b_2]$. Then

$$AB = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix}.$$

If any of a_1, a_2, b_1 or b_2 equals zero then AB has an entire row or an entire column of zeros. A matrix with an entire row or column of zeros is not invertible. Thus assume a_1, a_2, b_1 and b_2 are non-zero. Now if we add $-a_2/a_1$ of the first row to the second row we get

$$\begin{bmatrix} a_1 b_1 & a_1 b_2 \\ 0 & 0 \end{bmatrix}.$$

Thus AB is not row-equivalent to the identity. Thus by Theorem 12 page 23, AB is not invertible.

Exercise 7: Let A be an $n \times n$ (square) matrix. Prove the following two statements:

(a) If A is invertible and $AB = 0$ for some $n \times n$ matrix B then $B = 0$.

(b) If A is not invertible, then there exists an $n \times n$ matrix B such that $AB = 0$ but $B \neq 0$.

Solution:

(a) $0 = A^{-1}0 = A^{-1}(AB) = (A^{-1}A)B = IB = B$. Thus $B = 0$.

(b) By Theorem 13 (ii) since A is not invertible $AX = 0$ must have a non-trivial solution v. Let B be the matrix all of whose columns are equal to v. Then $B \neq 0$ but $AB = 0$.

Exercise 8: Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Prove, using elementary row operations, that A is invertible if and only if $(ad - bc) \neq 0$.

Solution: Suppose

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{bmatrix}.$$

Then $AB = I$ implies the following system in u, r, s, t has a solution

$$\begin{aligned} au + bs &= 1 \\ cu + ds &= 0 \\ ar + bt &= 0 \\ cr + dt &= 1 \end{aligned}$$

because (x, y, z, w) is one such solution. The augmented coefficient matrix of this system is

$$\left[\begin{array}{cccc|c} a & 0 & b & 0 & 1 \\ c & 0 & d & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & c & 0 & d & 1 \end{array}\right]. \tag{16}$$

As long as $ad - bc \neq 0$ this system row-reduces to the following row-reduced echelon form

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & d/(ad - bc) \\ 0 & 1 & 0 & 0 & -b/(ad - bc) \\ 0 & 0 & 1 & 0 & -c/(ad - bc) \\ 0 & 0 & 0 & 1 & a/(ad - bc) \end{array}\right]$$

Thus we see that $x = d/(ad - bc)$, $y = -b/(ad - bc)$, $z = -c/(ad - bc)$, $w = a/(ad - bc)$ and

$$A^{-1} = \begin{bmatrix} d/(ad - bc) & -b/(ad - bc) \\ -c/(ad - bc) & a/(ad - bc) \end{bmatrix}.$$

Now it's a simple matter to check that

$$\begin{bmatrix} d/(ad - bc) & -b/(ad - bc) \\ -c/(ad - bc) & a/(ad - bc) \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now suppose that $ad - bc = 0$. We will show there is no solution. If $a = b = c = d = 0$ then obviously A has no inverse. So suppose WOLOG that $a \neq 0$ (because by elementary row and column operations we can move any of the four elements to be the top left entry, and elementary row and column operations do not change a matrix's status as being invertible or not). Subtracting $\frac{c}{a}$ times the 3rd row from the 4th row of (16) gives

$$\left[\begin{array}{cccc|c} a & 0 & b & 0 & 1 \\ c & 0 & d & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & c - \frac{c}{a}a & 0 & d - \frac{c}{a}b & 1 \end{array}\right].$$

Now $c - \frac{c}{a}a = 0$ and since $ad - bc = 0$ also $d - \frac{c}{a}b = 0$. Thus we get

$$\left[\begin{array}{cccc|c} a & 0 & b & 0 & 1 \\ c & 0 & d & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

and it follows that A is not invertible.

Exercise 9: An $n \times n$ matrix A is called **upper-triangular** if $a_{ij} = 0$ for $i > j$, that is, if every entry below the main diagonal is 0. Prove that an upper-triangular (square) matrix is invertible if and only if every entry on its main diagonal is different from zero.

Solution: Suppose that $a_{ii} \neq 0$ for all i. Then we can divide row i by a_{ii} to give a row-equivalent matrix which has all ones on the diagonal. Then by a sequence of elementary row operations we can turn all off diagonal elements into zeros. We can therefore row-reduce the matrix to be equivalent to the identity matrix. By Theorem 12 page 23, A is invertible.

Now suppose that some $a_{ii} = 0$. If *all* a_{ii}'s are zero then the last row of the matrix is all zeros. A matrix with a row of zeros cannot be row-equivalent to the identity so cannot be invertible. Thus we can assume there's at least one i such that $a_{ii} \neq 0$. Let i' be the largest such index, so that $a_{i'i'} = 0$ and $a_{ii} \neq 0$ for all $i > i'$. We can divide all rows with $i > i'$ by a_{ii} to give ones on the diagonal for those rows. We can then add multiples of those rows to row i' to turn row i' into an entire row of zeros. Since again A is row-equivalent to a matrix with an entire row of zeros, it cannot be invertible.

Exercise 10: Prove the following generalization of Exercise 6. If A is an $m \times n$ matrix and B is an $n \times m$ matrix and $n < m$, then AB is not invertible.

Solution: There are n colunms in A so the vector space generated by those columns has dimension no greater than n. All columns of AB are linear combinations of the columns of A. Thus the vector space generated by the columns of AB is contained in the vector space generated by the columns of A. Thus the column space of AB has dimension no greater than n. Thus the column space of the $m \times m$ matrix AB has dimension less or equal to n and $n < m$. Thus the columns of AB generate a space of dimension strictly less than m. Thus AB is not invertible.

Exercise 11: Let A be an $n \times m$ matrix. Show that by means of a finite number of elementary row and/or column operations one can pass from A to a matrix R which is both 'row-reduced echelon' and 'column-reduced echelon,' i.e., $R_{ij} = 0$ if $i \neq j$, $R_{ii} = 1$, $1 \leq i \leq r$, $R_{ii} = 0$ if $i > r$. Show that $R = PAQ$, where P is an invertible $m \times m$ matrix and Q is an invertible $n \times n$ matrix.

Solution: First put A in row-reduced echelon form, R'. So $\exists$ an invertible $m \times m$ matrix P such that $R' = PA$. Each row of R' is either all zeros or starts (on the left) with zeros, then has a one, then may have non-zero entries after the one. Suppose row i has a leading one in the j-th column. The j-th column has zeros in all other places except the i-th, so if we add a multiple of this *column* to another column then it only affects entries in the i-th row. Therefore a sequence of such operations can turn this row into a row of all zeros and a single one.

Let B be the $n \times n$ matrix such that $B_{rr} = 1$ and $B_{rs} = 0 \ \forall \ r \neq s$ except $B_{jk} \neq 0$. Then AB equals A with B_{jk} times column j added to column k. B is invertible since any such operation can be undone by another such operation. By a sequence of such operations we can turn all values *after* the leading one into zeros. Let Q be a product of all of the elementary matrices B involved in this transformation. Then PAQ is in row-reduced and column-reduced form.

Exercise 12: The result of Example 16 suggests that perhaps the matrix

$$\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

is invertible and A^{-1} has integer entries. Can you prove that?

Solution: This problem seems a bit hard for this book. There are a class of theorems like this, in particular these are called Hilbert Matrices and a proof is given in this article on arxiv by Christian Berg called *Fibonacci numbers and orthogonal polynomials* (`http://arxiv.org/pdf/math/0609283v2.pdf`). See Theorem 4.1. Also there might be a more elementary proof in this discussion on mathoverflow.net where two proofs are given:
`http://mathoverflow.net/questions/47561/deriving-inverse-of-hilbert-matrix`.
Also see `http://vigo.ime.unicamp.br/HilbertMatrix.pdf` where a general formula for the i, j entry of the inverse is given explicitly as

$$(-1)^{i+j}(i+j-1)\binom{n+i-1}{n-j}\binom{n+j-1}{n-i}\binom{i+j-1}{i-1}^2$$

Chapter 2: Vector Spaces

Section 2.1: Vector Spaces

Page 29, three lines into the first paragraph of text they refer to "two operations". This could be confusing since they just said a vector space is a composite object consisting of a field and a set with a rule. There are two sets, the field F and the set of vectors V. Really there are four operations going around. There is addition and multiplication in the field F, and there is addition also in V, that's three. But there's also multiplication of an element of F and an element of V. That's why there are two distributive rules. But keep in mind we do not multiply elements of V together - there's no multiplication within V, only within F and between F and V. So anyway, why did they say "two" operations? They're clearly ignoring the two operations in the field and just talking about operations involving vectors. You get the point.

Exercise 1: If F is a field, verify that F^n (as defined in Example 1) is a vector space over the field F.

Solution: Example 1 starts with any field and defines the objects, the addition rule and the scalar multiplication rule. We must show the set of n-tuples satisfies the eight properties required in the definition.

1) Addition is commutative. Let $\alpha = (x_1, \dots, x_n)$ and $\beta = (y_1, \dots, y_n)$ be two n-tuples. Then $\alpha + \beta = (x_1 + y_1, \dots, x_n + y_n)$. And since F is commutative this equals $(y_1 + x_1, \dots, y_n + x_n)$, which equals $\beta + \alpha$. Thus $\alpha + \beta = \beta + \alpha$.

2) Addition is associative. Let $\alpha = (x_1, \dots, x_n)$, $\beta = (y_1, \dots, y_n)$ and $\gamma = (z_1, \dots, z_n)$ be three n-tuples. Then $(\alpha + \beta) + \gamma = ((x_1+y_1)+z_1, \dots, (x_n+y_n)+z_n)$. And since F is associative this equals $(x_1+(y_1+z_1), \dots, x_n+(y_n+z_n))$, which equals $\alpha+(\beta+\gamma)$.

3) We must show there is a unique vector 0 in V such that $\alpha + 0 = \alpha\ \forall\ \alpha \in V$. Consider $(0_F, \dots, 0_F)$ the vector of all 0's of length n, where 0_F is the zero element of F. Then this vector satisfies the property that $(0_F, \dots, 0_F) + (x_1, \dots, x_n) = (0_F + x_1, \dots, 0_F + x_n) = (x_1, \dots, x_n)$ since $0_F + x = x\ \forall\ x \in F$. Thus $(0_F, \dots, 0_F) + \alpha = \alpha\ \forall \alpha \in V$. We must just show this vector is unique with respect to this property. Suppose $\beta = (x_1, \dots, x_n)$ also satisfies the property that $\beta + \alpha = \alpha$ for all $\alpha \in V$. Let $\alpha = (0_F, \dots, 0_F)$. Then $(x_1, \dots, x_n) = (x_1 + 0_F, \dots, x_n + 0_F) = (x_1, \dots, x_n) + (0_F, \dots, 0_F)$ and by definition of β this equals $(0_F, \dots, 0_F)$. Thus $(x_1, \dots, x_n) = (0_F, \dots, 0_F)$. Thus $\beta = \alpha$ and the zero element is unique.

4) We must show for each vector α there is a unique vector β such that $\alpha + \beta = 0$. Suppose $\alpha = (x_1, \dots, x_n)$. Let $\beta = (-x_1, \dots, -x_n)$. Then β has the required property $\alpha + \beta = 0$. We must show β is unique with respect to this property. Suppose also $\beta' = (x_1', \dots, x_n')$ also has this property. Then $\alpha+\beta = 0$ and $\alpha+\beta' = 0$. So $\beta = \beta+0 = \beta+(\alpha+\beta') = (\beta+\alpha)+\beta' = 0+\beta' = \beta'$.

5) Let 1_F be the multiplicative identity in F. Then $1_F \cdot (x_1, \dots, x_n) = (1 \cdot x_1, \dots, 1 \cdot x_n) = (x_1, \dots, x_n)$ since $1_F \cdot x = x\ \forall\ x \in F$. Thus $1_F\alpha = \alpha\ \forall\ \alpha \in V$.

6) Let $\alpha = (x_1, \dots, x_n)$. Then $(c_1c_2)\alpha = ((c_1c_2)x_1, \dots, (c_1c_2)x_n)$ and since multiplication in F is associative this equals $(c_1(c_2x_1), \dots, c_1(c_2x_n)) = c_1(c_2x_1, \dots c_2x_n) = c_1 \cdot (c_2\alpha)$.

7) Let $\alpha = (x_1, \dots, x_n)$ and $\beta = (y_1, \dots, y_n)$. Then $c(\alpha+\beta) = c(x_1+y_1, \dots, x_n+y_n) = (c(x_1+y_1), \dots, c(x_n+y_n))$ and since multiplication is distributive over addition in F this equals $(cx_1+cy_1, \dots, cx_n+xy_n)$. This then equals $(cx_1, \dots, cx_n)+(cy_1, \dots, cy_n) =$

$c(x_1, \ldots, x_n) + c(y_1, \ldots, y_n) = c\alpha + c\beta$. Thus $c(\alpha + \beta) = c\alpha + c\beta$.

8) Let $\alpha = (x_1, \ldots, x_n)$. Then $(c_1 + c_2)\alpha = ((c_1 + c_2)x_1, \ldots, (c_1 + c_2)x_n)$ and since multiplication distributes over addition in F this equals $(c_1x_1 + c_2x_1, \ldots, c_1x_n + c_2x_n) = (c_1x_1, \ldots, c_1x_n) + (c_2x_1, \ldots c_2x_n) = c_1(x_1, \ldots, x_n) + c_2(x_1, \ldots, x_n) = c_1\alpha + c_2\alpha$. Thus $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$.

Exercise 2: If V is a vector space over the field F, verify that

$$(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4) = [\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4$$

for all vectors $\alpha_1, \alpha_2, \alpha_3$ and α_4 in V.

Solution: This follows associativity and commutativity properties of V:

$$\begin{aligned}
&(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4)\\
&= (\alpha_2 + \alpha_1) + (\alpha_3 + \alpha_4)\\
&= \alpha_2 + [\alpha_1 + (\alpha_3 + \alpha_4)]\\
&= \alpha_2 + [(\alpha_1 + \alpha_3) + \alpha_4]\\
&= [\alpha_2 + (\alpha_1 + \alpha_3)] + \alpha_4\\
&= [\alpha_2 + (\alpha_3 + \alpha_1)] + \alpha_4.
\end{aligned}$$

Exercise 3: If C is the field of complex numbers, which vectors in $\mathbb{C}^3$ are linear combinations of $(1, 0, -1)$, $(0, 1, 1)$, and $(1, 1, 1)$?

Solution: If we make a matrix out of these three vectors

$$A = \left[\begin{array}{ccc} 1 & 0 & 1\\ 0 & 1 & 1\\ -1 & 1 & 1 \end{array}\right]$$

then if we row-reduce the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0\\ 0 & 1 & 1 & 0 & 1 & 0\\ -1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

we get

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1\\ 0 & 1 & 0 & -1 & 2 & -1\\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right].$$

Therefore the matrix is invertible and $AX = Y$ has a solution $X = A^{-1}Y$ for any Y. Thus any vector $Y \in \mathbb{C}^3$ can be written as a linear combination of the three vectors. Not sure what the point was of making the base field $\mathbb{C}$.

Exercise 4: Let V be the set of all pairs (x, y) of real numbers, and let F be the field of real numbers. Define

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$$

$$c(x, y) = (cx, y).$$

Is V, with these operations, a vector space over the field of real numbers?

Solution: No it is not a vector space because $(0,2) = (0,1) + (0,1) = 2(0,1) = (2 \cdot 0, 1) = (0,1)$. Thus we must have $(0,2) = (0,1)$ which implies $1 = 2$ which is a contradiction in the field of real numbers.

Exercise 5: On $\mathbb{R}^n$, define two operations

$$\alpha \oplus \beta = \alpha - \beta$$

$$c \cdot \alpha = -c\alpha.$$

The operations on the right are the usual ones. Which of the axioms for a vector space are satisfied by $(\mathbb{R}^n, \oplus, \cdot)$?

Solution:

1) $\oplus$ is not commutative since $(0,\dots,0)\oplus(1,\dots,1) = (-1,\dots,-1)$ while $(1,\dots,1)\oplus(0,\dots,0) = (1,\dots,1)$. And $(1,\dots,1) \neq (-1,\dots,-1)$.

2) $\oplus$ is not associative since $((1,\dots,1)\oplus(1,\dots,1))\oplus(2,\dots,2) = (0,\dots,0)\oplus(2,\dots,2) = (-2,\dots,-2)$ while $(1,\dots,1)\oplus((1,\dots,1)\oplus(2,\dots,2)) = (1,\dots,1)\oplus(-1,\dots,-1) = (2,\dots,2)$.

3) There does exist a right additive identity, i.e. a vector 0 that satisfies $\alpha + 0 = \alpha$ for all α. The vector $\beta = (0,\dots,0)$ satisfies $\alpha + \beta = \alpha$ for all α. And if $\beta' = (b_1,\dots,b_n)$ also satisfies $(x_1,\dots,x_n) + \beta' = (x_1,\dots,x_n)$ then $x_i - b_i = x_i$ for all i and thus $b_i = 0$ for all i. Thus $\beta = (0,\dots,0)$ is unique with respect to the property $\alpha + \beta = \alpha$ for all α.

4) There do exist right additive inverses. For the vector $\alpha = (x_1,\dots,x_n)$ clearly only α itself satisfies $\alpha \oplus \alpha = (0,\dots,0)$.

5) The element 1 does not satisfy $1\cdot\alpha = \alpha$ for any non-zero α since otherwise we would have $1\cdot(x_1,\dots,x_n) = (-x_1,\dots,-x_n) = (x_1,\dots,x_n)$ only if $x_i = 0$ for all i.

6) The property $(c_1c_2)\cdot\alpha = c_1\cdot(c_2\cdot\alpha)$ does not hold since $(c_1c_2)\alpha = (-c_1c_2)\alpha$ while $c_1(c_2\alpha) = c_1(-c_2\alpha) = (-c_1(-c2\alpha)) = +c_1c_2\alpha$. Since $c_1c_2 \neq -c_1c_2$ for all c_1, c_2 they are not always equal.

7) It does hold that $c\cdot(\alpha\oplus\beta) = c\cdot\alpha \oplus c\cdot\beta$. Firstly, $c\cdot(\alpha\oplus\beta) = c\cdot(\alpha-\beta) = -c(\alpha-\beta) = -c\alpha + c\beta$. And secondly $c\cdot\alpha\oplus c\cdot\beta = (-c\alpha)\oplus(-c\beta) = -c\alpha - (-c\beta) = -c\alpha + c\beta$. Thus they are equal.

8) It does not hold that $(c_1 + c_2)\cdot\alpha = (c_1\cdot\alpha)\oplus(c_2\cdot\alpha)$. Firstly, $(c_1+c_2)\cdot\alpha = -(c_1+c_2)\alpha = -c_1\alpha - c_2\alpha$. Secondly, $c_1\cdot\alpha\oplus c_2\cdot\alpha = (-c_1\cdot\alpha)\oplus(-c_2\cdot\alpha) = -c_1\alpha + c_2\alpha$. Since $-c_1\alpha - c_2\alpha \neq -c_1\alpha - c_2\alpha$ for all c_1, c_2 they are not equal.

Exercise 6: Let V be the set of all complex-valued functions f on the real line such that (for all t in $\mathbb{R}$)

$$f(-t) = \overline{f(t)}.$$

The bar denotes complex conjugation. Show that V, with the operations

$$(f+g)(t) = f(t) + g(t)$$

$$(cf)(t) = cf(t)$$

is a vector space over the field of real numbers. Give an example of a function in V which is not real-valued.

Solution: We will use the basic fact that $\overline{a+b} = \overline{a} + \overline{b}$ and $\overline{ab} = \overline{a}\cdot\overline{b}$.

Before we show V satisfies the eight properties we must first show vector addition and scalar multiplication as defined are actually well-defined in the sense that they are indeed operations on V. In other words if f and g are two functions in V then we must show that $f+g$ is in V. In other words if $f(-t) = \overline{f(t)}$ and $g(-t) = \overline{g(t)}$ then we must show that $(f+g)(-t) = \overline{(f+g)(t)}$.

This is true because $(f+g)(-t) = f(-t)+g(-t) = \overline{f(t)} + \overline{g(t)} = \overline{(f(t)+g(t)} = \overline{(f+g)(t)}$.

Similary, if $c \in \mathbb{R}$, $(cf)(-t) = cf(-t) = c\overline{f(t)} = \overline{cf(t)}$ since $\overline{c} = c$ when $c \in \mathbb{R}$.

Thus the operations are well defined. We now show the eight properties hold:

1) Addition on functions in V is defined by adding in $\mathbb{C}$ to the values of the functions in $\mathbb{C}$. Thus since $\mathbb{C}$ is commutative, addition in V inherits this commutativity.

2) Similar to 1, since $\mathbb{C}$ is associative, addition in V inherits this associativity.

3) The zero function $g(t) = 0$ is in V since $-0 = \overline{0}$. And g satisfies $f+g = f$ for all $f \in V$. Thus V has a right additive identity.

4) Let g be the function $g(t) = -f(t)$. Then $g(-t) = -f(-t) = -\overline{f(t)} = \overline{-f(t)} = \overline{g(t)}$. Thus $g \in V$. And $(f+g)(t) = f(t)+g(t) = f(t) - f(t) = 0$. Thus g is a right additive inverse for f.

5) Clearly $1 \cdot f = f$ since 1 is the multiplicative identity in $\mathbb{R}$.

6) As before, associativity in $\mathbb{C}$ implies $(c_1c_2)f = c_1(c_2f)$.

7) Similarly, the distributive property in $\mathbb{C}$ implies $c(f+g) = cf + cg$.

8) Similarly, the distributive property in $\mathbb{C}$ implies $(c_1+c_2)f = c_1f + c_2f$.

An example of a function in V which is not real valued is $f(x) = ix$. Since $f(1) = i$ f is not real-valued. And $f(-x) = -ix = \overline{ix}$ since $x \in \mathbb{R}$, so $f \in V$.

Exercise 7: Let V be the set of pairs (x, y) of real numbers and let F be the field of real numbers. Define

$$(x, y) + (x_1, y_1) = (x + x_1, 0)$$

$$c(x, y) = (cx, 0).$$

Is V, with these operations, a vector space?

Solution: This is not a vector space because there would have to be an additive identity element (a, b) which has the property that $(a, b) + (x, y) = (x, y)$ for all $(x, y) \in V$. But this is impossible, because $(a, b) + (0, 1) = (a, 0) \neq (0, 1)$ no matter what (a, b) is. Thus V does not satisfy the third requirement of having an additive identity element.

Section 2.2: Subspaces

Page 39, typo in Exercise 3. It says $\mathbb{R}^5$, should be $\mathbb{R}^4$.

Exercise 1: Which of the following sets of vectors $\alpha = (a_1, \dots, a_n)$ in $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$ $(n \geq 3)$?

(a) all α such that $a_1 \geq 0$;

(b) all α such that $a_1 + 3a_2 = a_3$;

(c) all α such that $a_2 = a_1^2$;

(d) all α such that $a_1a_2 = 0$;

(e) all α such that a_2 is rational.

Solution:

(a) This is not a subspace because for $(1, \dots, 1)$ the additive inverse is $(-1, \dots, -1)$ which does not satisfy the condition.

(b) Suppose $(a_1, a_2, a_3, \dots, a_n)$ and $(b_1, b_2, b_3, \dots, b_n)$ satisfy the condition and let $c \in \mathbb{R}$. By Theorem 1 (page 35) we must show that $c(a_1, a_2, a_3, \dots, a_n)+(b_1, b_2, b_3, \dots, b_n) = (ca_1+b_1, \dots, ca_n+b_n)$ satisfies the condition. Now $(ca_1+b_1)+3(ca_2+b_2) = c(a_1 + 3a_2) + (b_1 + 3b_2) = c(a_3) + (b_3) = ca_3 + b_3$. Thus it does satisfy the condition so V is a vector space.

(c) This is not a vector space because $(1, 1)$ satisfies the condition since $1^2 = 1$, but $(1, 1, \dots) + (1, 1, \dots) = (2, 2, \dots)$ and $(2, 2, \dots)$ does not satisfy the condition because $2^2 \neq 2$.

(d) This is not a subspace. $(1, 0, \dots)$ and $(0, 1, \dots)$ both satisfy the condition, but their sum is $(1, 1, \dots)$ which does not satisfy the condition.

(e) This is not a subspace. $(1, 1, \dots, 1)$ satisfies the condition, but $\pi(1, 1, \dots, 1) = (\pi, \pi, \dots, \pi)$ does not satisfy the condition.

Exercise 2: Let V be the (real) vector space of all functions f from $\mathbb{R}$ into $\mathbb{R}$. Which of the following sets of functions are subspaces of V?

(a) all f such that $f(x^2) = f(x)^2$;

(b) all f such that $f(0) = f(1)$;

(c) all f such that $f(3) = 1 + f(-5)$;

(d) all f such that $f(-1) = 0$;

(e) all f which are continuous.

Solution:

(a) Not a subspace. Let $f(x) = x$ and $g(x) = x^2$. Then both satisfy the condition: $f(x^2) = x^2 = (f(x))^2$ and $g(x^2) = (x^2)^2 = (g(x))^2$. But $(f + g)(x) = x + x^2$ and $(f + g)(x^2) = x^2 + x^4$ while $[(f + g)(x)]^2 = (x + x^2)^2 = x^4 + 2x^3 + x^2$. These are not equal polynomials so the condition does not hold for $f + g$.

(b) Yes a subspace. Suppose f and g satisfy the property. Let $c \in \mathbb{R}$. Then $(cf+g)(0) = cf(0)+g(0) = cf(1)+g(1) = (cf+g)(1)$. Thus $(cf + g)(0) = (cf + g)(1)$. By Theorem 1 (page 35) the set of all such functions constitute a subspace.

(c) Not a subspae. Let $f(x)$ be the function defined by $f(3) = 1$ and $f(x) = 0$ for all $x \neq 3$. Let $g(x)$ be the function defined by $g(-5) = 0$ and $g(x) = 1$ for all $x \neq -5$. Then both f and g satisfy the condition. But $(f + g)(3) = f(3) + g(3) = 1 + 1 = 2$, while $1 + (f + g)(-5) = 1 + f(-5) + g(-5) = 1 + 0 + 0 = 1$. Since $1 \neq 2$, $f + g$ does not satisfy the condition.

(d) Yes a subspace. Suppose f and g satisfy the property. Let $c \in \mathbb{R}$. Then $(cf + g)(-1) = cf(-1) + g(-1) = c \cdot 0 + 0 = 0$. Thus $(cf + g)(-1) = 0$. By Theorem 1 (page 35) the set of all such functions constitute a subspace.

(e) Yes a subspace. Let f and g be continuous functions from $\mathbb{R}$ to $\mathbb{R}$ and let $c \in \mathbb{R}$. Then we know from basic results of real analysis that the sum and product of continuous functions are continuous. Since the function $c \mapsto c$ is continuous as well as f and g, it follows that $cf+g$ is continuous. By Theorem 1 (page 35) the set of all cotinuous functions constitute a subspace.

Exercise 3: Is the vector $(3, -1, 0, -1)$ in the subspace of $\mathbb{R}^5$ (sic) spanned by the vectors $(2, -1, 3, 2)$, $(-1, 1, 1, -3)$, and $(1, 1, 9, -5)$?

Solution: I assume they meant $\mathbb{R}^4$. No, $(3,-1,0,-1)$ is not in the subspace. If we row reduce the augmented matrix

$$\left[\begin{array}{ccc|c} 2 & -1 & 1 & 3 \\ -1 & 1 & 1 & -1 \\ 3 & 1 & 9 & 0 \\ 2 & -3 & -5 & -1 \end{array}\right]$$

we obtain

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 2 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & -2 \end{array}\right].$$

The two bottom rows are zero rows to the left of the divider, but the values to the right of the divider in those two rows are non-zero. Thus the system does not have a solution (see comments bottom of page 24 and top of page 25).

Exercise 4: Let W be the set of all (x_1,x_2,x_3,x_4,x_5) in $\mathbb{R}^5$ which satisfy

$$2x_1 - x_2 + \frac{4}{3}x_3 - x_4 \qquad = 0$$

$$x_1 \quad + \frac{2}{3}x_3 \quad - x_5 = 0$$

$$9x_1 - 3x_2 + 6x_3 - 3x_4 - 3x_5 = 0.$$

Find a finite set of vectors which spans W.

Solution: The matrix of the system is

$$\left[\begin{array}{ccccc} 2 & -1 & 4/3 & -1 & 0 \\ 1 & 0 & 2/3 & 0 & -1 \\ 9 & -3 & 6 & -3 & -3 \end{array}\right].$$

Row reducing to reduced echelon form gives

$$\left[\begin{array}{ccccc} 1 & 0 & 2/3 & 0 & -1 \\ 0 & 1 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Thus the system is equivalent to

$$x_1 + 2/3x_3 - x_5 = 0$$

$$x_2 + x_4 - 2x_5 = 0.$$

Thus the system is parametrized by (x_3,x_4,x_5). Setting each equal to one and the other two to zero (as in Example 15, page 42), in turn, gives the three vectors $(-2/3,0,1,0,0)$, $(0,-1,0,1,0)$ and $(1,2,0,0,1)$. These three vectors therefore span W.

Exercise 5: Let F be a field and let n be a positive integer ($n \geq 2$). Let V be the vector space of all $n \times n$ matrices over F. Which of the following sets of matrices A in V are subspaces of V?

(a) all invertible A;

(b) all non-invertible A;

(c) all A such that $AB = BA$, where B is some fixed matrix in V;

(d) all A such that $A^2 = A$.

Solution:

(a) This is not a subspace. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and let $B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Then both A and B are invertible, but $A+B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ which is not invertible. Thus the subset is not closed with respect to matrix addition. Therefore it cannot be a subspace.

(b) This is not a subspace. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and let $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Then neither A nor B is invertible, but $A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ which is invertible. Thus the subset is not closed with respect to matrix addition. Therefore it cannot be a subspace.

(c) This is a subspace. Suppose A_1 and A_2 satisfy $A_1B = BA_1$ and $A_2B = BA_2$. Let $c \in F$ be any constant. Then $(cA_1 + A_2)B = cA_1B + A_2B = cBA_1 + BA_2 = B(cA_1) + BA_2 = B(cA_1 + A_2)$. Thus $cA_1 + A_2$ satisfy the criteria. By Theorem 1 (page 35) the subset is a subspace.

(d) This is a subspace if F equals $\mathbb{Z}/2\mathbb{Z}$, otherwise it is not a subspace.
Suppose first that char$(F) \neq 2$. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $A^2 = A$. But $A + A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^2 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \neq \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Thus $A + A$ does not satisfy the criteria so the subset cannot be a subspace.

Suppose now that $F = \mathbb{Z}/2\mathbb{Z}$. Suppose A and B both satisfy $A^2 = A$ and $B^2 = B$. Let $c \in F$ be any scalar. Then $(cA + B)^2 = c^2A^2 + 2cAB + B^2$. Now $2 = 0$ so this reduces to $c^2A^2 + B^2$. If $c = 0$ then this reduces to B^2 which equals B, if $c = 1$ then this reduces to $A^2 + B^2$ which equals $A + B$. In both cases $(cA + B)^2 = cA + B$. Thus in this case by Theorem 1 (page 35) the subset is a subspace.

Finally suppose char$(F) = 2$ but F is not $\mathbb{Z}/2\mathbb{Z}$. Then $|F| > 2$. The polynomial $x^2 - x = 0$ has at most two solutions in F, so $\exists\, c \in F$ such that $c^2 \neq c$. Consider the identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $I^2 = I$. If such matrices form a subspace then it must be that cI is also in the subspace. Thus it must be that $(cI)^2 = cI$. Which is equivalent to $c^2 = c$, which contradicts the way c was chosen.

Exercise 6:

(a) Prove that the only subspaces of $\mathbb{R}^1$ are $\mathbb{R}^1$ and the zero subspace.

(b) Prove that a subspace of $\mathbb{R}^2$ is $\mathbb{R}^2$, or the zero subspace, or consists of all scalar multiples of some fixed vector in $\mathbb{R}^2$. (The last type of subspace is, intuitively, a straight line through the origin.)

(c) Can you describe the subspaces of $\mathbb{R}^3$?

Solution:

(a) Let V be a subspace of $\mathbb{R}^1$. Suppose $\exists\, v \in V$ with $v \neq 0$. Then v is a vector but it is also simply an element of $\mathbb{R}$. Let $\alpha \in \mathbb{R}$. Then $\alpha = \frac{\alpha}{v} \cdot v$ where $\frac{\alpha}{v}$ is a scalar in the base field $\mathbb{R}$. Since $cv \in V\ \forall\, c \in \mathbb{R}$, it follows that $\alpha \in V$. Thus we have shown that if $V \neq \{0\}$ then $V = \mathbb{R}^1$.

(b) We know the subsests $\{(0, 0)\}$ (example 6a, page 35) and $\mathbb{R}^2$ (example 1, page 29) are subspaces of $\mathbb{R}^2$. Also for any vector v in any vector space over any field F, the set $\{cv \mid c \in F\}$ is a subspace (Theorem 3, page 37). Thus we will be done if we show that if V is a subspace of $\mathbb{R}^2$ and there exists $v_1, v_2 \in V$ such that v_1 and v_2 do not lie on the same line, then $V = \mathbb{R}^2$. Equivalently we must show that any vector $w \in \mathbb{R}^2$ can be written as a linear combination of v_1 and v_2 whenever v_1 and v_2 are not co-linear. Equivalently, by Theorem 13 (iii) (page 23), it suffices to show that if $v_1 = (a, b)$ and $v_2 = (c, d)$ are not colinear, then the matrix $A = [v_1^T\ \ v_2^T]$ is invertible. Suppose $a \neq 0$ and let $x = c/a$. Then $xa = c$, and since v_1 and v_2 are not colinear, it follows that $xb \neq d$. Thus equivalently $ad - bc \neq 0$. It follows now from Exercise 1.6.8 pae 27 that if v_1 and v_2 not colinear then the matrix A^T is invertible. Finally A^T is invertible implies A is invertible, since clearly $(A^T)^{-1} = (A^{-1})^T$. Similarly if

$a = 0$ then it must be that $b \neq 0$ so we can make the same argument. So in all cases A is invertible.

(c) The subspaces are the zero subspace $\{0, 0, 0\}$, lines $\{cv \mid c \in \mathbb{R}\}$ for fixed $v \in \mathbb{R}^3$, planes $\{c_1v_1 + c_2v_2 \mid c_1, c_2 \in \mathbb{R}\}$ for fixed $v_1, v_2 \in \mathbb{R}^3$ and the whole space $\mathbb{R}^3$. By Theorem 3 we know these all are subspaces, we just must show they are the only subspaces. It suffices to show that if v_1, v_2 and v_3 are not co-planar then the space generated by v_1, v_2, v_3 is all of $\mathbb{R}^3$. Equivalently we must show if v_1 and v_2 are not co-linear, and v_3 is not in the plane generated by v_1, v_2 then any vector $w \in \mathbb{R}^3$ can be written as a linear combination of v_1, v_2, v_3. Equivalently, by Theorem 13 (iii) (page 23), it suffices to show the matrix $A = [v_1\ \ v_2\ \ v_3]$ is invertible. A is invertible $\Leftrightarrow A^T$ is invertible, since clearly $(A^T)^{-1} = (A^{-1})^T$. Now v_3 is in the plane generated by $v_1, v_2 \Leftrightarrow v_3$ can be written as a linear combination of v_1 and $v_2 \Leftrightarrow A^T$ is row equivalent to a matrix with one of its rows equal to all zeros (this follows from Theorem 12, page 23) $\Leftrightarrow A^T$ is *not* invertible. Thus v_3 is not in the plane generated by $v_1, v_2 \Leftrightarrow A$ is invertible.

Exercise 7: Let W_1 and W_2 be subspaces of a vector space V such that the set-theoretic union of W_1 and W_2 is also a subspace. Prove that one of the spaces W_i is contained in the other.

Solution: Assume the space generated by W_1 and W_2 is equal to their set-theoretic union $W_1 \cup W_2$. Suppose $W_1 \not\subseteq W_2$ and $W_2 \not\subseteq W_1$. We wish to derive a contradiction. So suppose $\exists\, w_1 \in W_1 \setminus W_2$ and $\exists\, w_2 \in W_2 \setminus W_1$. Consider $w_1 + w_2$. By assumption this is in $W_1 \cup W_2$, so $\exists\, w_1' \in W_1$ such that $w_1 + w_2 = w_1'$ or $\exists\, w_2' \in W_2$ such that $w_1 + w_2 = w_2'$. If the former, then $w_2 = w_1' - w_1 \in W_1$ which contradicts the assumption that $w_2 \notin W_1$. Likewise the latter implies the contradiction $w_1 \in W_2$. Thus we are done.

Exercise 8: Let V be the vector space of all functions from $\mathbb{R}$ into $\mathbb{R}$; let V_e be the susbset of even functions, $f(-x) = f(x)$; let V_o be the subset of odd functions, $f(-x) = -f(x)$.

(a) Prove that V_e and V_o are subspaces of V.

(b) Prove that $V_e + V_o = V$.

(c) Prove that $V_e \cap V_o = \{0\}$.

Solution:

(a) Let $f, g \in V_e$ and $c \in \mathbb{R}$. Let $h = cf + g$. Then $h(-x) = cf(-x) + g(-x) = cf(x) + g(x) = h(x)$. So $h \in V_e$. By Theorem 1 (page 35) V_e is a subspace. Now let $f, g \in V_o$ and $c \in \mathbb{R}$. Let $h = cf+g$. Then $h(-x) = cf(-x)+g(-x) = -cf(x)-g(x) = -h(x)$. So $h \in V_o$. By Theorem 1 (page 35) V_o is a subspace.

(b) Let $f \in V$. Let $f_e(x) = \frac{f(x)+f(-x)}{2}$ and $f_o = \frac{f(x)-f(-x)}{2}$. Then f_e is even and f_o is odd and $f = f_e + f_o$. Thus we have written f as $g + h$ where $g \in V_e$ and $h \in V_o$.

(c) Let $f \in V_e \cap V_o$. Then $f(-x) = -f(x)$ and $f(-x) = f(x)$. Thus $f(x) = -f(x)$ which implies $2f(x) = 0$ which implies $f = 0$.

Exercise 9: Let W_1 and W_2 be subspaces of a vector space V such that $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$. Prove that for each vector α in V there are *unique* vectors α_1 in W_1 and α_2 in W_2 such that $\alpha = \alpha_1 + \alpha_2$.

Solution: Let $\alpha \in W$. Suppose $\alpha = \alpha_1 + \alpha_2$ for $\alpha_i \in W_i$ and $\alpha = \beta_1 + \beta_2$ for $\beta_i \in W_i$. Then $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$ which implies $\alpha_1 - \beta_1 = \beta_2 - \alpha_2$. Thus $\alpha_1 - \beta_1 \in W_1$ and $\alpha_1 - \beta_1 \in W_2$. Since $W_1 \cap W_2 = \{0\}$ it follows that $\alpha_1 - \beta_1 = 0$ and thus $\alpha_1 = \beta_1$. Similarly, $\beta_2 - \alpha_2 \in W_1 \cap W_2$ so also $\alpha_2 = \beta_2$.

Section 2.3: Bases and Dimension

Exercise 1: Prove that if two vectors are linearly dependent, one of them is a scalar multiple of the other.

Solution: Suppose v_1 and v_2 are linearly dependent. If one of them, say v_1, is the zero vector then it is a scalar multiple of the other one $v_1 = 0 \cdot v_2$. So we can assume both v_1 and v_2 are non-zero. Then if $\exists\, c_1, c_2$ such that $c_1v_1 + c_2v_2 = 0$, both c_1 and c_2 must be non-zero. Therefore we can write $v_1 = -\frac{c_2}{c_1}v_2$.

Exercise 2: Are the vectors

$$\alpha_1 = (1,1,2,4), \quad \alpha_2 = (2,-1,-5,2)$$
$$\alpha_3 = (1,-1,-4,0), \quad \alpha_4 = (2,1,1,6)$$

linearly independent in $\mathbb{R}^4$?

Solution: By Corollary 3, page 46, it suffices to determine if the matrix whose rows are the α_i's is invertible. By Theorem 12 (ii) we can do this by row reducing the matrix

$$\begin{bmatrix} 1 & 1 & 2 & 4 \\ 2 & -1 & -5 & 2 \\ 1 & -1 & -4 & 0 \\ 2 & 1 & 1 & 6 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 1 & 2 & 4 \\ 2 & -1 & -5 & 2 \\ 1 & -1 & -4 & 0 \\ 2 & 1 & 1 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & -3 & -9 & -6 \\ 0 & -2 & -6 & -4 \\ 0 & -1 & -3 & -2 \end{bmatrix} \underset{\text{rows}}{\overset{\text{swap}}{\to}} \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & -1 & -3 & -2 \\ 0 & -3 & -9 & -6 \\ 0 & -2 & -6 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 3 & 2 \\ 0 & -3 & -9 & -6 \\ 0 & -2 & -6 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus the four vectors are not linearly independent.

Exercise 3: Find a basis for the subspace of $\mathbb{R}^4$ spanned by the four vectors of Exercise 2.

Solution: In Section 2.5, Theorem 9, page 56, it will be proven that row equivalent matrices have the same row space. The proof of this is almost immediate so there seems no easier way to prove it than to use that fact. If you multiply a matrix A on the left by another matrix P, the rows of the new matrix PA are linear combinations of the rows of the original matrix. Thus the rows of PA generate a subspace of the space generated by the rows of A. If P is invertible, then the two spaces must be contained in each other since we can go backwards with P^{-1}. Thus the rows of row-equivalent matrices generate the same space. Thus using the row reduced form of the matrix in Exercise 2, it must be that the space is two dimensoinal and generated by $(1,1,2,4)$ and $(0,1,3,2)$.

Exercise 4: Show that the vectors

$$\alpha_1 = (1,0,-1), \quad \alpha_2 = (1,2,1), \quad \alpha_3 = (0,-3,2)$$

form a basis for $\mathbb{R}^3$. Express each of the standard basis vectors as linear combinations of α_1, α_2, and α_3.

Solution: By Corollary 3, page 46, to show the vectors are linearly independent it suffices to show the matrix whose rows are the α_i's is invertible. By Theorem 12 (ii) we can do this by row reducing the matrix

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 0 & -3 & 2 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 0 & -3 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 2 \\ 0 & -3 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & -3 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now to write the standard basis vectors in terms of these vectors, by the discussion at the bottom of page 25 through page 26, we can row-reduce the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & -3 & 2 & 0 & 0 & 1 \end{array}\right].$$

This gives

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & -3 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & 2 & -1 & 1 & 0 \\ 0 & -3 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1/2 & 1/2 & 0 \\ 0 & -3 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1/2 & 1/2 & 0 \\ 0 & 0 & 5 & -3/2 & 3/2 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1/2 & 1/2 & 0 \\ 0 & 0 & 1 & -3/10 & 3/10 & 1/5 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7/10 & 3/10 & 1/5 \\ 0 & 1 & 0 & -1/5 & 1/5 & -1/5 \\ 0 & 0 & 1 & -3/10 & 3/10 & 1/5 \end{array}\right].$$

Thus if

$$P = \left[\begin{array}{ccc} 7/10 & 3/10 & 1/5 \\ -1/5 & 1/5 & -1/5 \\ -3/10 & 3/10 & 1/5 \end{array}\right]$$

then $PA = I$, so we have

$$\frac{7}{10}\alpha_1 + \frac{3}{10}\alpha_2 + \frac{1}{5}\alpha_3 = (1, 0, 0)$$

$$-\frac{1}{5}\alpha_1 + \frac{1}{5}\alpha_2 - \frac{1}{5}\alpha_3 = (0, 1, 0)$$

$$-\frac{3}{10}\alpha_1 + \frac{3}{10}\alpha_2 + \frac{1}{5}\alpha_3 = (0, 0, 1).$$

Exercise 5: Find three vectors in $\mathbb{R}^3$ which are linearly dependent, and are such that any two of them are linearly independent.

Solution: Let $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ and $v_3 = (1, 1, 0)$. Then $v_1 + v_2 - v_3 = (0, 0, 0)$ so they are linearly dependent. We know v_1 and v_2 are linearly independent as they are two of the standard basis vectors (see Example 13, page 41). Suppose $av_1 + bv_3 = 0$. Then $(a + b, b, 0) = (0, 0, 0)$. The second coordinate implies $b = 0$ and then the first coordinate in turn implies $a = 0$. Thus v_1 and v_3 are linearly independent. Analogously v_2 and v_3 are linearly independent.

Exercise 6: Let V be the vector space of all 2×2 matrices over the field F. Prove that V has dimension 4 by exhibiting a basis for V which has four elements.

Solution: Let

$$v_{11} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right], \quad v_{12} = \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]$$

$$v_{21} = \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right], \quad v_{22} = \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$$

Suppose $av_{11} + bv_{12} + cv_{21} + dv_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

from which it follows immediately that $a = b = c = d = 0$. Thus $v_{11}, v_{12}, v_{21}, v_{22}$ are linearly independent.

Now let $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be any 2×2 matrix. Then $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = av_{11} + bv_{12} + cv_{21} + dv_{22}$. Thus $v_{11}, v_{12}, v_{21}, v_{22}$ span the space of 2×2 matrices.

Thus $v_{11}, v_{12}, v_{21}, v_{22}$ are both linearly independent and they span the space of all 2×2 matrices. Thus $v_{11}, v_{12}, v_{21}, v_{22}$ constitue a basis for the space of all 2×2 matrices.

Exercise 7: Let V be the vector space of Exercise 6. Let W_1 be the set of matrices of the form

$$\begin{bmatrix} x & -x \\ y & z \end{bmatrix}$$

and let W_2 be the set of matrices of the form

$$\begin{bmatrix} a & b \\ -a & c \end{bmatrix}.$$

(a) Prove that W_1 and W_2 are subspaces of V.

(b) Find the dimension of W_1, W_2, $W_1 + W_2$, and $W_1 \cap W_2$.

Solution:

(a) Let $A = \begin{bmatrix} x & -x \\ y & z \end{bmatrix}$ and $B = \begin{bmatrix} x' & -x' \\ y' & z' \end{bmatrix}$ be two elements of W_1 and let $c \in F$. Then

$$cA + B = \begin{bmatrix} cx + x' & -cx - x' \\ cy + y' & cz + z' \end{bmatrix} = \begin{bmatrix} a & -a \\ u & v \end{bmatrix}$$

where $a = cx + x'$, $u = cy + y'$ and $v = cz + z'$. Thus $cA + B$ is in the form of an element of W_1. Thus $cA + B \in W_1$. By Theorem 1 (page 35) W_1 is a subspace.

Now let $A = \begin{bmatrix} a & b \\ -a & d \end{bmatrix}$ and $B = \begin{bmatrix} a' & b' \\ -a' & d' \end{bmatrix}$ be two elements of W_1 and let $c \in F$.Then

$$cA + B = \begin{bmatrix} ca + a' & cb + b' \\ -ca - a' & cd + d' \end{bmatrix} = \begin{bmatrix} x & y \\ -x & z \end{bmatrix}$$

where $x = ca + a'$, $y = cb + b'$ and $z = cd + d'$. Thus $cA + B$ is in the form of an element of W_2. Thus $cA + B \in W_2$. By Theorem 1 (page 35) W_2 is a subspace.

(b) Let

$$A_1 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then $A_1, A_2, A_3 \in W_1$ and

$$c_1A_1 + c_2A_2 + c_3A_3 = \begin{bmatrix} c_1 & -c_1 \\ c_2 & c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

implies $c_1 = c_2 = c_3 = 0$. So A_1, A_2, A_3 are linearly independent. Now let $A = \begin{bmatrix} x & -x \\ y & z \end{bmatrix}$ be any element of W_1. Then $A = xA_1 + yA_2 + zA_3$. Thus A_1, A_2, A_3 span W_1. Thus $\{A_1, A_2, A_3\}$ form a basis for W_1. Thus W_1 has dimension three.

Let

$$A_1 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then $A_1, A_2, A_3 \in W_2$ and

$$c_1A_1 + c_2A_2 + c_3A_3 = \begin{bmatrix} c_1 & c_2 \\ -c_1 & c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

implies $c_1 = c_2 = c_3 = 0$. So A_1, A_2, A_3 are linearly independent. Now let $A = \begin{bmatrix} x & y \\ -x & z \end{bmatrix}$ be any element of W_2. Then $A = xA_1 + yA_2 + zA_3$. Thus A_1, A_2, A_3 span W_2. Thus $\{A_1, A_2, A_3\}$ form a basis for W_2. Thus W_2 has dimension three.

Let V be the space of 2×2 matrices. We showed in Exercise 6 that the $\dim(V) = 4$. Now $W_1 \subseteq W_1 + W_2 \subseteq V$. Thus by Corollary 1, page 46, $3 \leq \dim(W_1 + W_2) \leq 4$. Let $A = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$. Then $A \in W_2$ and $A \notin W_1$. Thus $W_1 + W_2$ is strictly bigger than W_1. Thus $4 \geq \dim(W_1 + W_2) > \dim(W_1) = 3$. Thus $\dim(W_1 + W_2) = 4$.

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is in $W_1 \cap W_2$. Then $A \in W_1 \Rightarrow a = -b$ and $A \in W_2 \Rightarrow a = -c$. So $A = \begin{bmatrix} a & -a \\ -a & b \end{bmatrix}$. Let $A_1 = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Suppose $aA_1 + bA_2 = 0$. Then

$$\begin{bmatrix} a & -a \\ -a & b \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which implies $a = b = 0$. Thus A_1 and A_2 are linearly independent. Let $A = \begin{bmatrix} a & -a \\ -a & b \end{bmatrix} \in W_1 \cap W_2$. Then $A = aA_1 + bA_2$. So A_1, A_2 span $W_1 \cap W_2$. Thus $\{A_1, A_2\}$ is a basis for $W_1 \cap W_2$. Thus $\dim(W_1 \cap W_2) = 2$.

Exercise 8: Again let V be the space of 2×2 matrices over F. Find a basis $\{A_1, A_2, A_3, A_4\}$ for V such that $A_j^2 = A_j$ for each j.

Solution: Let V be the space of all 2×2 matrices. Let

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_4 \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

Then $A_i^2 = A_i \; \forall \, i$. Now

$$aA_1 + bA_2 + cA_3 + dA_4 = \begin{bmatrix} a+c & c \\ d & b+d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

implies $c = d = 0$ which in turn implies $a = b = 0$. Thus A_1, A_2, A_3, A_4 are linearly independent. Thus they span a subspace of A of dimension four. But by Exercise 6, A also has dimension four. Thus by Corollary 1, page 46, the subspace spanned by A_1, A_2, A_3, A_4 is the entire space. Thus $\{A_1, A_2, A_3, A_4\}$ is a basis.

Exercise 9: Let V be a vector space over a subfield F of the complex numbers. Suppose α, β, and γ are linearly independent vectors in V. Prove that $(\alpha + \beta)$, $(\beta + \gamma)$, and $(\gamma + \alpha)$ are linearly independent.

Solution: Suppose $a(\alpha+\beta)+b(\beta+\gamma)+c(\gamma+\alpha)=0$. Rearranging gives $(a+c)\alpha+(a+b)\beta+(b+c)\gamma=0$. Since α,β, and γ are linearly independent it follows that $a+c=a+b=b+c=0$. This gives a system of equations in a,b,c with matrix

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

This row-reduces as follows:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since this row-reduces to the identity matrix, by Theorem 7, page 13, the only solution is $a=b=c=0$. Thus $(\alpha+\beta)$, $(\beta+\gamma)$, and $(\gamma+\alpha)$ are linearly independent.

Exercise 10: Let V be a vector space over the field F. Suppose there are a finite number of vectors $\alpha_1,\dots,\alpha_r$ in V which span V. Prove that V is finite-dimensional.

Solution: If any α_i's are equal to zero then we can remove them from the set and the remaining α_i's still span V. Thus we can assume WLOG that $\alpha_i \neq 0\ \forall\ i$. If $\alpha_1,\dots,\alpha_r$ are linearly independent, then $\{\alpha_1,\dots,\alpha_r\}$ is a basis and $\dim(V)=r<\infty$. On the other hand if $\alpha_1,\dots,\alpha_r$ are linearly dependent, then $\exists\ c_1,\dots,c_r\in F$, not all zero, such that $c_1\alpha_1+\cdots+c_r\alpha_r=0$. Suppose WLOG that $c_r\neq 0$. Then $\alpha_r=-\frac{c_1}{c_r}\alpha_1-\cdots-\frac{c_{r-1}}{c_r}\alpha_{r-1}$. Thus α_r is in the subspace spanned by $\alpha_1,\dots,\alpha_{r-1}$. Thus $\alpha_1,\dots,\alpha_{r-1}$ spans V. If $\alpha_1,\dots,\alpha_{r-1}$ are linearly independent then $\{\alpha_1,\dots,\alpha_{r-1}\}$ is a basis and $\dim(V)=r-1<\infty$. If $\alpha_1,\dots,\alpha_{r-1}$ are linearly dependent then arguing as before (with possibly re-indexing) we can produce $\alpha_1,\dots,\alpha_{r-2}$ which span V. Continuing in this way we must eventually arrive at a linearly independent set, or arrive at a set that consists of a single element, that still spans V. If we arrive at a single element v_1 then $\{v_1\}$ is linearly independent since $cv_1=0\Rightarrow c=0$ (see comments after (2-9) page 31). Thus we must eventually arrive at a finite set that is spans and is linearly independent. Thus we must eventually arrive at a finite basis, which implies $\dim(V)<\infty$.

Exercise 11: Let V be the set of all 2×2 matrices A with *complex* entries which satisfy $A_{11}+A_{22}=0$.

(a) Show that V is a vector space over the field of *real* numbers, with the usual operations of matrix addition and multiplication of a matrix by a scalar.

(b) Find a basis for this vector space.

(c) Let W be the set of all matrices A in V such that $A_{21}=-\overline{A_{12}}$ (the bar denotes complex conjugation). Prove that W is a subspace of V and find a basis for W.

Solution: (a) It is clear from inspection of the definition of a vector space (pages 28-29) that a vector space over a field F is a vector space over every subfield of F, because all properties (e.g. commutativity and associativity) are inherited from the operations in F. Let M be the vector space of all 2×2 matrices over $\mathbb{C}$ (M *is* a vector space, see example 2 page 29). We will show V is a subspace M as a vector space over $\mathbb{C}$. It will follow from the comment above that V is a vector space over $\mathbb{R}$. Now V is a subset of M, so using Theorem 1 (page 35) we must show whenever $A,B\in V$ and $c\in\mathbb{C}$ then $cA+B\in V$. Let $A,B\in V$. Write $A=\begin{bmatrix} x & y \\ z & w \end{bmatrix}$ and $B=\begin{bmatrix} x' & y' \\ z' & w' \end{bmatrix}$. Then

$$x+w=x'+w'=0. \tag{17}$$

$$cA+B=\begin{bmatrix} cx+x' & cy+y' \\ cz+z' & cw+w' \end{bmatrix}$$

To show $cA+B\in V$ we must show $(cx+x')+(cw+w')=0$. Rearranging the left hand side gives $c(x+w)+(x'+w')$ which equals zero by (17).

(b) We can write the general element of V as

$$A = \begin{bmatrix} a+bi & e+fi \\ g+hi & -a-bi \end{bmatrix}.$$

Let

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},$$
$$v_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 & i \\ 0 & 0 \end{bmatrix},$$
$$v_5 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad v_6 = \begin{bmatrix} 0 & 0 \\ i & 0 \end{bmatrix}.$$

Then $A = av_1 + bv_2 + ev_3 + fv_4 + gv_5 + hv_6$ so $v_1, v_2, v_3, v_4, v_5, v_6$ span V. Suppose $av_1 + bv_2 + ev_3 + fv_4 + gv_5 + hv_6 = 0$. Then

$$av_1 + bv_2 + ev_3 + fv_4 + gv_5 + hv_6 = \begin{bmatrix} a+bi & e+fi \\ g+hi & -a-bi \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

implies $a = b = c = d = e = f = g = h = 0$ because a complex number $u + vi = 0 \Leftrightarrow u = v = 0$. Thus $v_1, v_2, v_3, v_4, v_5, v_6$ are linearly independent. Thus $\{v_1, \ldots, v_6\}$ is a basis for V as a vector space over $\mathbb{R}$, and $\dim(V) = 6$.

(c) Let $A, B \in W$ and $c \in \mathbb{R}$. By Theorem 1 (page 35) we must show $cA + B \in W$. Write $A = \begin{bmatrix} x & y \\ -\bar{y} & -x \end{bmatrix}$ and $B = \begin{bmatrix} x' & y' \\ -\bar{y}' & -x' \end{bmatrix}$, where $x, y, x', y' \in \mathbb{C}$. Then

$$cA + B = \begin{bmatrix} cx + x' & cy + y' \\ -c\bar{y} - \bar{y}' & -cx - x' \end{bmatrix}.$$

Since $-c\bar{y} - \bar{y}' = -\overline{(cy + y')}$, it follows that $cA + B \in W$. Note that we definitely need $c \in \mathbb{R}$ for this to be true.

It remains to find a basis for W. We can write the general element of W as

$$A = \begin{bmatrix} a+bi & e+fi \\ -e+fi & -a-bi \end{bmatrix}.$$

Let

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},$$
$$v_3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Then $A = av_1 + bv_2 + ev_3 + fv_4$ so v_1, v_2, v_3, v_4 span V. Suppose $av_1 + bv_2 + ev_3 + fv_4 = 0$. Then

$$av_1 + bv_2 + ev_3 + fv_4 = \begin{bmatrix} a+bi & e+fi \\ -e+fi & -a-bi \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

implies $a = b = e = f = 0$ because a complex number $u + vi = 0 \Leftrightarrow u = v = 0$. Thus v_1, v_2, v_3, v_4 are linearly independent. Thus $\{v_1, \ldots, v_4\}$ is a basis for V as a vector space over $\mathbb{R}$, and $\dim(V) = 4$.

Exercise 12: Prove that the space of $m \times n$ matrices over the field F has dimension mn, by exhibiting a basis for this space.

Solution: Let M be the space of all $m \times n$ matrices. Let M_{ij} be the matrix of all zeros except for the i, j-th place which is a one. We claim $\{M_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ constitute a basis for M. Let $A = (a_{ij})$ be an arbitrary marrix in M. Then

$A = \sum_{ij} a_{ij} M_{ij}$. Thus $\{M_{ij}\}$ span M. Suppose $\sum_{ij} a_{ij} M_{ij} = 0$. The left hand side equals the matrix (a_{ij}) and this equals the zero matrix if and only if every $a_{ij} = 0$. Thus $\{M_{ij}\}$ are linearly independent as well. Thus the *nm* matrices constitute a basis and M has dimension *mn*.

Exercise 13: Discuss Exercise 9, when V is a vector space over the field with two elements described in Exercise 5, Section 1.1.

Solution: If F has characteristic two then $(\alpha+\beta)+(\beta+\gamma)+(\gamma+\alpha) = 2\alpha+2\beta+2\gamma = 0+0+0 = 0$ since in a field of characteristic two, $2 = 0$. Thus in this case $(\alpha+\beta)$, $(\beta+\gamma)$ and $(\gamma+\alpha)$ are linearly dependent. However any two of them are linearly independent. For example suppose $a_1(\alpha+\beta)+a_2(\beta+\gamma) = 0$. The LHS equals $a_1\alpha + a_2\gamma + (a_1+a_2)\beta$. Since α, β, γ are linearly independent, this is zero only if $a_1 = 0$, $a_2 = 0$ and $a_1 + a_2 = 0$. In particular $a_1 = a_2 = 0$, so $\alpha+\beta$ and $\beta+\gamma$ are linearly independent.

Exercise 14: Let V be the set of real numbers. Regard V as a vector space over the field of *rational* numbers, with the usual operations. Prove that this vector space is *not* finite-dimensional.

Solution: We know that $\mathbb{Q}$ is countable and $\mathbb{R}$ is uncountable. Since the set of n-tuples of things from a countable set is countable, $\mathbb{Q}^n$ is countable for all n. Now, suppose $\{r_1, \dots, r_n\}$ is a basis for $\mathbb{R}$ over $\mathbb{Q}$. Then every element of $\mathbb{R}$ can be written as $a_1 r_1 + \cdots + a_n r_n$. Thus we can map n-tuples of rational numbers onto $\mathbb{R}$ by $(a_1, \dots, a_n) \mapsto a_1 r_1 + \cdots + a_n r_n$. Thus the cardinality of $\mathbb{R}$ must be less or equal to $\mathbb{Q}^n$. But the former is uncountable and the latter is countable, a contradiction. Thus there can be no such finite basis.

Section 2.4: Coordinates

Exercise 1: Show that the vectors

$$\alpha_1 = (1,1,0,0), \quad \alpha_2 = (0,0,1,1)$$

$$\alpha_3 = (1,0,0,4), \quad \alpha_4 = (0,0,0,2)$$

form a basis for $\mathbb{R}^4$. Find the coordinates of each of the standard basis vectors in the ordered basis $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$.

Solution: Using Theorem 7, page 52, if we calculate the inverse of

$$P = \left[\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 4 & 2 \end{array}\right].$$

then the columns of P^{-1} will give the coefficients to write the standard basis vectors in terms of the α_i's. We do this by row-reducing the augmented matrix

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 4 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

The left side must reduce to the identity whlie the right side transforms to the inverse of P. Row reduction gives

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 4 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 4 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 4 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 & 0 & -1 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 & 0 & -1 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & -4 & 4 & -1 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 2 & -1/2 & 1/2 \end{array}\right].$$

Thus $\{\alpha_1, \dots, \alpha_4\}$ is a basis. Call this basis β. Thus $(1,0,0,0) = \alpha_3 - 2\alpha_4$, $(0,1,0,0) = \alpha_1 - \alpha_3 + 2\alpha_4$, $(0,0,1,0) = \alpha_2 - \frac{1}{2}\alpha_4$ and $(0,0,0,1) = \frac{1}{2}\alpha_4$.

Thus $[(1,0,0,0)]_\beta = (0,0,1,-2)$, $[(0,1,0,0)]_\beta = (1,0,-1,2)$, $[(0,0,1,0)]_\beta = (0,1,0,-1/2)$ and $[(0,0,0,1)]_\beta = (0,0,0,1/2)$.

Exercise 2: Find the coordinate matrix of the vector $(1,0,1)$ in the basis of $\mathbb{C}^3$ consisting of the vectors $(2i,1,0)$, $(2,-1,0)$, $(0,1+i,1-i)$, in that order.

Solution: Using Theorem 7, page 52, the answer is $P^{-1}\left[\begin{array}{c} 1 \\ 0 \\ 1 \end{array}\right]$ where

$$P = \left[\begin{array}{ccc} 2i & 2 & 0 \\ 1 & -1 & 1+i \\ 0 & 0 & 1-i \end{array}\right].$$

We find P^{-1} by row-reducing the augmented matrix

$$\left[\begin{array}{ccc|ccc} 2i & 2 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1+i & 0 & 1 & 0 \\ 0 & 0 & 1-i & 0 & 0 & 1 \end{array}\right].$$

The right side will transform into the P^{-1}. Row reducing:

$$\left[\begin{array}{ccc|ccc} 2i & 2 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1+i & 0 & 1 & 0 \\ 0 & 0 & 1-i & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc} 1 & -1 & 1+i & 0 & 1 & 0 \\ 2i & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1-i & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc} 1 & -1 & 1+i & 0 & 1 & 0 \\ 0 & 2+2i & 2-2i & 1 & -2i & 0 \\ 0 & 0 & 1-i & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc} 1 & -1 & 1+i & 0 & 1 & 0 \\ 0 & 1 & -i & \frac{1-i}{4} & \frac{-1-i}{2} & 0 \\ 0 & 0 & 1-i & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & \frac{1-i}{4} & \frac{1-i}{2} & 0 \\ 0 & 1 & -i & \frac{1-i}{4} & \frac{-1-i}{2} & 0 \\ 0 & 0 & 1-i & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & \frac{1-i}{4} & \frac{1-i}{2} & 0 \\ 0 & 1 & -i & \frac{1-i}{4} & \frac{-1-i}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1+i}{2} \end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1-i}{4} & \frac{1-i}{2} & \frac{-1-i}{2} \\ 0 & 1 & 0 & \frac{1-i}{4} & \frac{-1-i}{2} & \frac{-1+i}{2} \\ 0 & 0 & 1 & 0 & 0 & \frac{1+i}{2} \end{array}\right]$$

Therefore

$$P^{-1}\left[\begin{array}{c} 1 \\ 0 \\ 1 \end{array}\right]=\left[\begin{array}{ccc} \frac{1-i}{4} & \frac{1-i}{2} & \frac{-1-i}{2} \\ \frac{1-i}{4} & \frac{-1-i}{2} & \frac{-1+i}{2} \\ 0 & 0 & \frac{1+i}{2} \end{array}\right]\left[\begin{array}{c} 1 \\ 0 \\ 1 \end{array}\right]=\left[\begin{array}{c} \frac{-1-3i}{4} \\ \frac{-1+i}{4} \\ \frac{1+i}{2} \end{array}\right]$$

Thus $(1,0,1)=\frac{-1-3i}{4}(2i,1,0)+\frac{-1+i}{4}(2,-1,0)+\frac{1+i}{2}(0,1+i,1-i)$.

Exercise 3: Let $\mathcal{B}=\{\alpha_1,\alpha_2,\alpha_3\}$ be the ordered basis for $\mathbb{R}^3$ consisting of

$$\alpha_1=(1,0,-1),\quad \alpha_2=(1,1,1),\quad \alpha_3=(1,0,0).$$

What are the coordinates of the vector (a,b,c) in the ordered basis $\mathcal{B}$?

Solution: Using Theorem 7, page 52, the answer is $P^{-1}\left[\begin{array}{c} a \\ b \\ c \end{array}\right]$ where

$$P=\left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{array}\right].$$

We find P^{-1} by row-reducing the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

The right side will transform into the P^{-1}. Row reducing:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right].$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right].$$

Therefore,

$$P^{-1}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b-c \\ b \\ a-2b+c \end{bmatrix}$$

Thus the answer is

$$[(a,b,c)]_{\mathcal{B}} = (b-c,\ b,\ a-2b+c).$$

Exercise 4: Let W be the subspace of $\mathbb{C}^3$ spanned by $\alpha_1 = (1,\ 0,\ i)$ and $\alpha_2 = (1+i,\ 1,\ -1)$.

(a) Show that α_1 and α_2 form a basis for W.

(b) Show that the vectors $\beta_1 = (1,\ 1,\ 0)$ and $\beta_2 = (1,\ i,\ 1+i)$ are in W and form another basis for W.

(c) What are the coordinates of α_1 and α_2 in the ordered basis $\{\beta_1, \beta_2\}$ for W?

Solution: (a) To show α_1 and α_2 form a basis of the space they generate we must show they are linearly independent. In other words that $a\alpha_1 + b\alpha_2 = 0 \Rightarrow a = b = 0$. Equivalently we need to show neither is a multiple of the other. If $\alpha_2 = c\alpha_1$ then from the second coordinate it follows that $c = 0$ which would imply $\alpha_2 = (0,0,0)$, which it does not. So $\{\alpha_1, \alpha_2\}$ is a basis for the space they genearate.

(b) Since the first coordinate of both β_1 and β_2 is one, it's clear that neither is a multiple of the other. So they generate a two dimensional subspace of $\mathbb{C}^3$. If we show β_1 and β_2 can be written as linear combinations of α_1 and α_2 then since the spaces generated by them both have dimension two, by Corollary 1, page 46, they must be equal. To show β_1 and β_2 can be written as linear combinations of α_1 and α_2 we row-reduce the augmented matrix

$$\left[\begin{array}{cc|cc} 1 & 1+i & 1 & 1 \\ 0 & 1 & 1 & i \\ i & -1 & 0 & 1+i \end{array}\right].$$

Row reduction follows:

$$\left[\begin{array}{cc|cc} 1 & 1+i & 1 & 1 \\ 0 & 1 & 1 & i \\ i & -1 & 0 & 1+i \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1+i & 1 & 1 \\ 0 & 1 & 1 & i \\ 0 & -i & -i & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -i & 2-i \\ 0 & 1 & 1 & i \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\beta_1 = -i\alpha_1 + \alpha_2$ and $\beta_2 = (2-i)\alpha_1 + i\alpha_2$.

(c) We have to write the β_i's in terms of the α_i's, basically the opposite of what we did in part b. In this case we row-reduce the augmented matrix

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1+i \\ 1 & i & 0 & 1 \\ 0 & 1+i & i & -1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1+i \\ 0 & -1+i & -1 & -i \\ 0 & 1+i & i & -1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1+i \\ 0 & -1+i & -1 & -i \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\to\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1+i \\ 0 & 1 & \frac{1+i}{2} & \frac{-1+i}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]\to\left[\begin{array}{cc|cc} 1 & 0 & \frac{1-i}{2} & \frac{3+i}{2} \\ 0 & 1 & \frac{1+i}{2} & \frac{-1+i}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\alpha_1 = \frac{1-i}{2}\beta_1 + \frac{1+i}{2}\beta_2$ and $\alpha_2 = \frac{3+i}{2}\beta_1 + \frac{-1+i}{2}\beta_2$. So finally, if $\mathcal{B}$ is the basis $\{\beta_1, \beta_2\}$ then

$$[\alpha_1]_{\mathcal{B}} = \left(\frac{1-i}{2}, \frac{1+i}{2}\right)$$

$$[\alpha_2]_{\mathcal{B}} = \left(\frac{3+i}{2}, \frac{-1+i}{2}\right).$$

Exercise 5: Let $\alpha = (x_1, x_2)$ and $\beta = (y_1, y_2)$ be vectors in $\mathbb{R}^2$ such that

$$x_1y_1 + x_2y_2 = 0, \quad x_1^2 + x_2^2 = y_1^2 + y_2^2 = 1.$$

Prove that $\mathcal{B} = \{\alpha, \beta\}$ is a basis for $\mathbb{R}^2$. Find the coordinates of the vector (a, b) in the ordered basis $\mathcal{B} = \{\alpha, \beta\}$. (The conditions on α and β say, geometrically, that α and β are perpendicular and each has length 1.)

Solution: It suffices by Corollary 1, page 46, to show α and β are linearly indepdenent, because then they generate a subspace of $\mathbb{R}^2$ of dimension two, which therefore must be all of $\mathbb{R}^2$. The second condition on x_1, x_2, y_1, y_2 implies that neither α nor β are the zero vector. To show two vectors are linearly independent we only need show neither is a non-zero scalar multiple of the other. Suppose WLOG that $\beta = c\alpha$ for some $c \in \mathbb{R}$, and since neither vector is the zero vector, $c \neq 0$. Then $y_1 = cx_1$ and $y_2 = cx_2$. Thus the conditions on x_1, x_2, y_1, y_2 implies

$$0 = x_1y_1 + x_2y_2 = cx_1^2 + cx_2^2 = c(x_1^2 + x_2^2) = c \cdot 1 = c.$$

Thus $c = 0$, a contradiction.

It remains to find the coordinates of the arbitrary vector (a, b) in the ordered basis $\{\alpha, \beta\}$. To find the coordinates of (a, b) we can row-reduce the augmented matrix

$$\left[\begin{array}{cc|c} x_1 & y_1 & a \\ x_2 & y_2 & b \end{array}\right].$$

It cannot be that both $x_1 = x_2 = 0$ so assume WLOG that $x_1 \neq 0$. Also it cannot be that both $y_1 = y_2 = 0$. Assume first that $y_1 \neq 0$. Since order matters we cannot assume $y_1 \neq 0$ WLOG, so we must consider both cases. Then note that $x_1y_1 + x_2y_2 = 0$ implies

$$\frac{x_2y_2}{x_1y_1} = -1 \tag{18}$$

Thus if $x_1y_2 - x_2y_1 = 0$ then $\frac{x_2}{x_1} = \frac{y_2}{y_1}$ from which (18) implies $\left(\frac{x_2}{x_1}\right)^2 = -1$, a contradiction. Thus we can conclude that $x_1y_2 - x_2y_1 \neq 0$. We use this in the following row reduction to be sure we are not dividing by zero.

$$\left[\begin{array}{cc|c} x_1 & y_1 & a \\ x_2 & y_2 & b \end{array}\right]\to\left[\begin{array}{cc|c} 1 & y_1/x_1 & a/x_1 \\ x_2 & y_2 & b \end{array}\right]\to\left[\begin{array}{cc|c} 1 & y_1/x_1 & a/x_1 \\ 0 & y_2 - \frac{x_2y_1}{x_1} & b - \frac{x_2a}{x_1} \end{array}\right]=\left[\begin{array}{cc|c} 1 & y_1/x_1 & a/x_1 \\ 0 & \frac{x_1y_2 - x_2y_1}{x_1} & \frac{bx_1 - ax_2}{x_1} \end{array}\right]$$

$$\to\left[\begin{array}{cc|c} 1 & y_1/x_1 & a/x_1 \\ 0 & 1 & \frac{bx_1 - ax_2}{x_1y_2 - x_2y_1} \end{array}\right]\to\left[\begin{array}{cc|c} 1 & 0 & \frac{ay_2 - by_1}{x_1y_2 - x_2y_1} \\ 0 & 1 & \frac{bx_1 - ax_2}{x_1y_2 - x_2y_1} \end{array}\right]$$

Now if we substitute $y_1 = -x_2y_2/x_1$ into the numerator and denominator of $\frac{ay_2 - by_1}{x_1y_2 - x_2y_1}$ and use $x_1^2 + x_2^2 = 1$ it simplifies to $ax_1 + bx_2$. Similarly $\frac{ay_2 - by_1}{x_1y_2 - x_2y_1}$ simplifies to $ay_1 + by_2$. So we get

$$\left[\begin{array}{cc|c} 1 & 0 & ax_1 + bx_2 \\ 0 & 1 & ay_1 + by_2 \end{array}\right].$$

Now assume $y_2 \neq 0$ (and we continue to assume $x_1 \neq 0$ since we assumed that WLOG). In this case

$$\frac{y_1}{y_2} = -\frac{x_2}{x_1} \tag{19}$$

So if $x_1y_2 - x_2y_1 = 0$ then $\frac{x_2y_1}{x_1y_2} = 1$. But then (19) implies $\left(\frac{x_2}{x_1}\right)^2 = -1$ a contradition. So also in this case we can assume $x_1y_2 - x_2y_1 \neq 0$ and so we can do the same row-reduction as before. Thus in all cases

$$(ax_1 + bx_2)\alpha + (ay_1 + by_2)\beta = (a, b)$$

or equivalently

$$(ax_1 + bx_2)(x_1, x_2) + (ay_1 + by_2)(y_1, y_2) = (a, b).$$

Exercise 6: Let V be the vector space over the complex numbers of all functions from $\mathbb{R}$ into $\mathbb{C}$, i.e., the space of all complex-valued functions on the real line. Let $f_1(x) = 1$, $f_2(x) = e^{ix}$, $f_3(x) = e^{-ix}$.

(a) Prove that f_1, f_2, and f_3 are linearly independent.

(b) Let $g_1(x) = 1$, $g_2(x) = \cos x$, $g_3(x) = \sin x$. Find an invertible 3×3 matrix P such that

$$g_j = \sum_{i=1}^{3} P_{ij} f_i.$$

Solution: Suppose $a + be^{ix} + ce^{-ix} = 0$ as functions of $x \in \mathbb{R}$. In other words $a + be^{ix} + ce^{-ix} = 0$ for all $x \in \mathbb{R}$. Let $y = e^{ix}$. Then $y \neq 0$ and $a + by + \frac{c}{y} = 0$ which implies $ay + by^2 + c = 0$. This is at most a quadratic polynomial in y thus can be zero for at most two values of y. But e^{ix} takes infinitely many different values as x varies in $\mathbb{R}$, so $ay + by^2 + c$ cannot be zero for all $y = e^{ix}$, so this is a contradiction.

We know that $e^{ix} = \cos(x) + i\sin(x)$. Thus $e^{-ix} = \cos(x) - i\sin(x)$. Adding these gives $2\cos(x) = e^{ix} + e^{-ix}$. Thus $\cos(x) = \frac{1}{2}e^{ix} + \frac{1}{2}e^{-ix}$. Subtracting instead of adding the equations gives $e^{ix} - e^{-ix} = 2i\sin(x)$. Thus $\sin(x) = \frac{1}{2i}e^{ix} - \frac{1}{2i}e^{-ix}$ or equivalently $\sin(x) = -\frac{i}{2}e^{ix} + \frac{i}{2}e^{-ix}$. Thus the requested matrix is

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -i/2 \\ 0 & 1/2 & i/2 \end{bmatrix}.$$

Exercise 7: Let V be the (real) vector space of all polynomial functions from $\mathbb{R}$ into $\mathbb{R}$ of degree 2 or less, i.e., the space of all functions f of the form

$$f(x) = c_0 + c_1x + c_2x^2.$$

Let t be a fixed real number and define

$$g_1(x) = 1, \quad g_2(x) = x + t, \quad g_3(x) = (x + t)^2.$$

Prove that $\mathcal{B} = \{g_1, g_2, g_3\}$ is a basis for V. If

$$f(x) = c_0 + c_1x + c_2x^2$$

what are the coordinates of f in this ordered basis $\mathcal{B}$?

Solution: We know V has dimension three (it follows from Example 16, page 43, that $\{1, x, x^2\}$ is a basis). Thus by Corollary 2 (b), page 45, it suffices to show $\{g_1, g_2, g_3\}$ span V. We need to solve for u, v, w the equation

$$c_2x^2 + c_1x + c_0 = u + v(x + t) + w(x + t)^2.$$

Rearranging

$$c_2x^2 + c_1x + c_0 = wx^2 + (v + 2wt)x + (u + vt + wt^2).$$

It follows that

$$w = c_2$$
$$v = c_1 - 2c_2 t$$
$$u = c_0 - c_1 t + c_2 t^2.$$

Thus $\{g_1, g_2, g_3\}$ do span V and the coordinates of $f(x) = c_2x^2 + c_1x + c_0$ are

$$(c_2,\ \ c_1 - 2c_2 t,\ \ c_0 - c_1 t + c_2 t^2).$$

Section 2.6: Computations Concerning Subspaces

Exercise 1: Let $s < n$ and A an $s \times n$ matrix with entries in the field F. Use Theorem 4 (not its proof) to show that there is a non-zero X in $F^{n\times 1}$ such that $AX = 0$.

Solution: Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the colunms of A. Then $\alpha_i \in F^{s\times 1}\ \forall\ i$. Thus $\{\alpha_1, \ldots, \alpha_n\}$ are n vectors in $F^{s\times 1}$. But $F^{s\times 1}$ has dimension $s < n$ thus by Theorem 4, page 44, $\alpha_1, \ldots, \alpha_n$ cannot be linearly independent. Thus $\exists\ x_1, \ldots, x_n \in F$ such that $x_1\alpha_1 + \cdots + x_n\alpha_n = 0$. Thus if

$$X = \left[\begin{array}{c} x_1 \\ \vdots \\ x_n \end{array}\right]$$

then $AX = x_1\alpha_1 + \cdots + x_n\alpha_n = 0$.

Exercise 2: Let

$$\alpha_1 = (1, 1, -2, 1), \quad \alpha_2 = (3, 0, 4, -1), \quad \alpha_3 = (-1, 2, 5, 2).$$

Let

$$\alpha = (4, -5, 9, -7), \quad \beta = (3, 1, -4, 4), \quad \gamma = (-1, 1, 0, 1).$$

(a) Which of the vectors α, β, γ are in the subspace of $\mathbb{R}^4$ spanned by the α_i?

(b) Which of the vectors α, β, γ are in the subspace of $\mathbb{C}^4$ spanned by the α_i?

(c) Does this suggest a theorem?

Solution:

(a) We use the approach of row-reducing the matrix whose rows are given by the α_i:

$$\left[\begin{array}{cccc} 1 & 1 & -2 & 1 \\ 3 & 0 & 4 & -1 \\ -1 & 2 & 5 & 2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccc} 1 & 1 & -2 & 1 \\ 0 & -3 & 10 & -4 \\ 0 & 3 & 3 & 3 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccc} 1 & 1 & -2 & 1 \\ 0 & 0 & 13 & -1 \\ 0 & 1 & 1 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccc} 1 & 0 & -3 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 13 & -1 \end{array}\right]$$

$$\rightarrow \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1/13 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & -3/13 \\ 0 & 1 & 0 & 14/13 \\ 0 & 0 & 1 & -1/13 \end{bmatrix}$$

Let $\rho_1 = (1, 0, 0, -3/13)$, $\rho_2 = (0, 1, 0, 14/13)$ and $\rho_3 = (0, 0, 1, -1/13)$. Thus elements of the subspace spanned by the α_i are of the form $b_1\rho_1 + b_2\rho_2 + b_3\rho_3$

$$= \left(b_1,\ b_2,\ b_3,\ \tfrac{1}{13}(14b_2 - 3b_1 - b_3)\right).$$

- $\alpha = (4, -5, 9, -7)$. We have $b_1 = 4$, $b_2 = -5$ and $b_3 = 9$. Thus if α is in the subspace it must be that

$$\frac{1}{13}(14(-5) - 3(4) - 9) \stackrel{?}{=} b_4$$

  where $b_4 = -7$. Indeed the left hand side does equal -7, so α is in the subspace.

- $\beta = (3, 1, -4, 4)$. We have $b_1 = 3$, $b_2 = 1$, $b_3 = -4$. Thus if β is in the subspace it must be that

$$\frac{1}{13}(14 - 3(3) + 4) \stackrel{?}{=} b_4$$

  where $b_4 = 4$. But the left hand side equals $9/13 \neq 4$ so β is *not* in the subspace.

- $\gamma = (-1, 1, 0, 1)$. We have $b_1 = -1$, $b_2 = 1$, $b_3 = 0$. Thus if γ is in the subspace it must be that

$$\frac{1}{13}(14 - 3(-1) - 0) \stackrel{?}{=} b_4$$

  where $b_4 = 1$. But the left hand side equals $17/13 \neq 1$ so γ is *not* in the subspace.

(b) Nowhere in the above did we use the fact that the field was $\mathbb{R}$ instead of $\mathbb{C}$. The only equations we had to solve are linear equations with real coefficients, which have solutions in $\mathbb{R}$ if and only if they have solutions in $\mathbb{C}$. Thus the same results hold: α is in the subspace while β and γ are not.

(c) This suggests the following theorem: Suppose F is a subfield of the field E and $\alpha_1, \ldots, \alpha_n$ are a basis for a subspace of F^n, and $\alpha \in F^n$. Then α is in the subspace of F^n generated by $\alpha_1, \ldots, \alpha_n$ if and only if α is in the subspace of E^n generated by $\alpha_1, \ldots, \alpha_n$.

Exercise 3: Consider the vectors in $\mathbb{R}^4$ defined by

$$\alpha_1 = (-1, 0, 1, 2), \quad \alpha_2 = (3, 4, -2, 5), \quad \alpha_3 = (1, 4, 0, 9).$$

Find a system of homogeneous linear equations for which the space of solutions is exactly the subspace of $\mathbb{R}^4$ spanned by the three given vectors.

Solution: We use the approach of row-reducing the matrix whose rows are given by the α_i:

$$\begin{bmatrix} -1 & 0 & 1 & 2 \\ 3 & 4 & -2 & 5 \\ 1 & 4 & 0 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 4 & 1 & 11 \\ 0 & 4 & 1 & 11 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1/4 & 11/4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let $\rho_1 = (1, 0, -1, -2)$ and $\rho_2 = (0, 1, 1/4, 11/4)$. Then the arbitrary element of the subspace spanned by α_1 and α_2 is of the form $b_1\rho_1 + b_2\rho_2$ for arbitrary $b_1, b_2 \in \mathbb{R}$. Expanding we get

$$b_1\rho_1 + b_2\rho_2 = (b_1,\ \ b_2,\ \ -b_1 + \frac{1}{4}b_2,\ \ -2b_2 + \frac{11}{4}b_2).$$

Thus the equations that must be satisfied for (x, y, z, w) to be in the subspace are

$$\left\{\begin{array}{l} z = -x + \frac{1}{4}y \\ w = -2x + \frac{11}{4}y \end{array}\right. .$$

or equivalently

$$\left\{\begin{array}{l} -x + \frac{1}{4}y - z = 0 \\ -2x + \frac{11}{4}y - w = 0 \end{array}\right. .$$

Exercise 4: In $\mathbb{C}^3$ let

$$\alpha_1 = (1, 0, -i), \quad \alpha_2 = (1+i, 1-i, 1), \quad \alpha_3 = (i, i, i).$$

Prove that these vectors form a basis for $\mathbb{C}^3$. What are the coordinates of the vector (a, b, c) in this basis?

Solution: We use the approach of row-reducing the augmented matrix:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -i & 1 & 0 & 0 \\ 1+i & 1-i & 1 & 0 & 1 & 0 \\ i & i & i & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -i & 1 & 0 & 0 \\ 0 & 1-i & i & -1-i & 1 & 0 \\ 0 & i & i-1 & -i & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -i & 1 & 0 & 0 \\ 0 & 1 & \frac{-1+i}{2} & -i & \frac{1+i}{2} & 0 \\ 0 & i & i-1 & -i & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -i & 1 & 0 & 0 \\ 0 & 1 & \frac{-1+i}{2} & -i & \frac{1+i}{2} & 0 \\ 0 & 0 & \frac{-1+3i}{2} & -1-i & \frac{1-i}{2} & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -i & 1 & 0 & 0 \\ 0 & 1 & \frac{-1+i}{2} & -i & \frac{1+i}{2} & 0 \\ 0 & 0 & 1 & \frac{-2+4i}{5} & \frac{-2-i}{5} & \frac{-1-3i}{5} \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1-2i}{5} & \frac{1-2i}{5} & \frac{3-i}{5} \\ 0 & 1 & 0 & \frac{1-2i}{5} & \frac{1+3i}{5} & \frac{-2-i}{5} \\ 0 & 0 & 1 & \frac{-2+4i}{5} & \frac{-2-i}{5} & \frac{-1-3i}{5} \end{array}\right]$$

Since the left side transformed into the identity matrix we know that $\{\alpha_1, \alpha_2, \alpha_3\}$ form a basis for $\mathbb{C}^3$. We used the vectors to form the rows of the augmented matrix not the columns, so the matrix on the right is $(P^{\mathrm{T}})^{-1}$ from (2-17). But $(P^{\mathrm{T}})^{-1} = (P^{-1})^{\mathrm{T}}$, so the coordinate matrix of (a, b, c) with respect to the basis $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ are given by

$$[(a, b, c)]_{\mathcal{B}} = (P^{-1})^{\mathrm{T}} \left[\begin{array}{c} a \\ b \\ c \end{array}\right]$$

$$= \left[\begin{array}{c} \frac{1-2i}{5}a + \frac{1-2i}{5}b + \frac{-2+4i}{5}c \\ \frac{1-2i}{5}a + \frac{1+3i}{5}b + \frac{-2-i}{5}c \\ \frac{3-i}{5}a + \frac{-2-i}{5}b + \frac{-1-3i}{5}c \end{array}\right].$$

Exercise 5: Give an explicit description of the type (2-25) for the vectors

$$\beta = (b_1, b_2, b_3, b_4, b_5)$$

in $\mathbb{R}^5$ which are linear combinations of the vectors

$$\alpha_1 = (1,0,2,1,-1), \quad \alpha_2 = (-1,2,-4,2,0)$$

$$\alpha_3 = (2,-1,5,2,1), \quad \alpha_4 = (2,1,3,5,2).$$

Solution: We row-reduce the matrix whose rows are given by the α_i's.

$$\left[\begin{array}{ccccc} 1 & 0 & 2 & 1 & -1 \\ -1 & 2 & -4 & 2 & 0 \\ 2 & -1 & 5 & 2 & 1 \\ 2 & 1 & 3 & 5 & 2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 0 & 2 & 1 & -1 \\ 0 & 2 & -2 & 3 & -1 \\ 0 & -1 & 1 & 0 & 3 \\ 0 & 1 & -1 & 3 & 4 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 0 & 2 & 1 & -1 \\ 0 & 1 & -1 & 3 & 4 \\ 0 & 0 & 0 & -3 & -9 \\ 0 & 0 & 0 & 3 & 7 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 0 & 2 & 1 & -1 \\ 0 & 1 & -1 & 3 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 3 & 7 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 0 & 2 & 1 & -4 \\ 0 & 1 & -1 & 0 & -5 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 0 & 2 & 0 & -4 \\ 0 & 1 & -1 & 0 & -5 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Let $\rho_1 = (1,0,2,0,0)$, $\rho_2 = (0,1,-1,0,0)$, $\rho_3 = (0,0,0,1,0)$ and $\rho_4 = (0,0,0,0,1)$. Then the general element that is a linear combination of the α_i's is $b_1\rho_1 + b_2\rho_2 + b_3\rho_3 + b_4\rho_4 = (b_1,\ b_2,\ 2b_1 - b_2,\ b_3,\ b_4)$.

Exercise 6: Let V be the real vector space spanned by the rows of the matrix

$$A = \left[\begin{array}{ccccc} 3 & 21 & 0 & 9 & 0 \\ 1 & 7 & -1 & -2 & -1 \\ 2 & 14 & 0 & 6 & 1 \\ 6 & 42 & -1 & 13 & 0 \end{array}\right].$$

(a) Find a basis for V.

(b) Tell which vectors $(x_1, x_2, x_3, x_4, x_5)$ are elements of V.

(c) If $(x_1, x_2, x_3, x_4, x_5)$ is in V what are its coordinates in the basis chosen in part (a)?

Solution: We row-reduce the matrix

$$\left[\begin{array}{ccccc} 3 & 21 & 0 & 9 & 0 \\ 1 & 7 & -1 & -2 & -1 \\ 2 & 14 & 0 & 6 & 1 \\ 6 & 42 & -1 & 13 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 7 & -1 & -2 & -1 \\ 0 & 0 & 3 & 15 & 3 \\ 0 & 0 & 2 & 10 & 3 \\ 0 & 0 & 5 & 25 & 6 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 7 & -1 & -2 & -1 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 2 & 10 & 3 \\ 0 & 0 & 5 & 25 & 6 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 7 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc} 1 & 7 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

(a) A basis for V is given by the non-zero rows of the reduced matrix

$$\rho_1 = (1, 7, 0, 3, 0), \quad \rho_2 = (0, 0, 1, 5, 0), \quad \rho_3 = (0, 0, 0, 0, 1).$$

(b) Vectors of V are any of the form $b_1\rho_1 + b_2\rho_2 + b_3\rho_3$

$$= (b_1,\ 7b_1,\ b_2,\ 3b_1 + 5b_2,\ b_3)$$

for arbitrary $b_1, b_2, b_3 \in \mathbb{R}$.

(c) By the above, the element $(x_1, x_2, x_3, x_4, x_5)$ in V must be of the form $x_1\rho_1 + x_3\rho_2 + x_5\rho_3$. In other words if $\mathcal{B} = \{\rho_1, \rho_2, \rho_3\}$ is the basis for V given in part (a), then the coordinate matrix of $(x_1, x_2, x_3, x_4, x_5)$ is

$$[(x_1, x_2, x_3, x_4, x_5)]_{\mathcal{B}} = \left[\begin{array}{c} x_1 \\ x_3 \\ x_5 \end{array}\right].$$

Exercise 7: Let A be an $m \times n$ matrix over the field F, and consider the system of equations $AX = Y$. Prove that this system of equations has a solution if and only if the row rank of A is equal to the row rank of the augmented matrix of the system.

Solution: To solve the system we row-reduce the augmented matrix $[A\,|\,Y]$ resulting in an augmented matrix $[R\,|\,Z]$ where R is in reduced echelon form and Z is an $m \times 1$ matrix. If the last k rows of R are zero rows then the system has a solution if and only if the last k entries of Z are also zeros. Thus the only non-zero entries in Z are in the non-zero rows of R. These rows are already linearly independent, and they clearly remain independent regardless of the augmented values. Thus if there are solutions then the rank of the augmented matrix is the same as the rank of R. Conversely, if there are non-zero entries in Z in any of the last k rows then the system has no solutions. We want to show that those non-zero rows in the augmented matrix are linearly independent from the non-zero rows of R, so we can conclude that the rank of R is less than the rank of $[R\,|\,Z]$. Let S

be the set of rows of $[R\,|\,Z]$ that contain all rows where R is non-zero, plus one additional row r where Z is non-zero. Suppose a linear combination of the elements of S equals zero. Since $c \cdot r = 0 \Leftrightarrow r = 0$, at least one of the elements of S different from r must have a non-zero coefficient. Suppose row $r' \in S$ has non-zero coefficient c in the linear combination. Suppose the leading one in row r' is in position i. Then the i-th coordinate of the linear combination is also c, because except for the one in the i-th position, all other entries in the i-th column of $[R\,|\,Z]$ are zero. Thus there can be no non-zero coefficients. Thus the set S is linearly independent and $|S| = |R| + 1$. Thus the system has a solution if and only if the rank of R is the same as the rank of $[R\,|\,Z]$. Now A has the same rank as R and $[R\,|\,Z]$ has the same rank as $[A\,|\,Y]$ since they differ by elementary row operations. Thus the system has a solution if and only if the rank of A is the same as the rank of $[A\,|\,Y]$.

Chapter 3: Linear Transformations

Section 3.1: Linear Transformations

Exercise 1: Which of the following functions T from $\mathbb{R}^2$ into $\mathbb{R}^2$ are linear transformations?

(a) $T(x_1, x_2) = (1 + x_1, x_2)$;

(b) $T(x_1, x_2) = (x_2, x_1)$;

(c) $T(x_1, x_2) = (x_1^2, x_2)$;

(d) $T(x_1, x_2) = (\sin x_1, x_2)$;

(e) $T(x_1, x_2) = (x_1 - x_2, 0)$.

Solution:

(a) T is not a linear transformation because $T(0, 0) = (1, 0)$ and according to the comments after Example 5 on page 68 we know that it must always be that $T(0, 0) = (0, 0)$.

(b) T is a linear transformation. Let $\alpha = (x_1, x_2)$ and $\beta = (y_1, y_2)$. Then $T(c\alpha+\beta) = T((cx_1+y_1, cx_2+y_2)) = (cx_2+y_2, cx_1+y_1) = c(x_2, x_1) + (y_2, y_1) = cT(\alpha) + T(\beta)$.

(c) T is not a linear transformation. If T were a linear transformation then we'd have $(1, 0) = T((-1, 0)) = T(-1 \cdot (1, 0)) = -1 \cdot T(1, 0) = -1 \cdot (1, 0) = (-1, 0)$ which is a contradiction, $(1, 0) \neq (-1, 0)$.

(d) T is not a linear transformation. If T were a linear transformation then $(0, 0) = T(\pi, 0) = T(2(\pi/2, 0)) = 2T((\pi/2, 0)) = 2(\sin(\pi/2), 0) = 2(1, 0) = (2, 0)$ which is a contradiction, $(0, 0) \neq (2, 0)$.

(e) T is a linear transformation. Let $Q = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$. Then (identifying $\mathbb{R}^2$ with $\mathbb{R}^{1\times 2}$) $T(x_1, x_2) = [x_1 \;\; x_2]Q$ so from Example 4, page 68, (with P being the identity matrix), it follows that T is a linear transformation.

Exercise 2: Find the range, rank, null space, and nullity for the zero transformation and the identity transformation on a finite-dimensional vector space V.

Solution: Suppose V has dimension n. The range of the zero transformation is the zero subspace $\{0\}$; the range of the identity transformation is the whole space V. The rank of the zero transformation is the dimension of the range which is zero; the rank of the identity transformation is the rank of the whole space V which is n. The null space of the zero transformation is the whole space V; the null space of the identity transformation is the zero subspace $\{0\}$. The nullity of the zero transformation is the dimension of its null space, which is the whole space, so is n; the nullity of the identity transformation is the dimension of its null space, which is the zero space, so is 0.

Exercise 3: Describe the range and the null space for the differentiation transformation of Example 2. Do the same for the integration transformation of Example 5.

Solution: V is the space of polynomals. The range of the differentiiation transformation is all of V since if $f(x) = c_0 + c_1x + \cdots + c_nx^n$ then $f(x) = (Dg)(x)$ where $g(x) = c_0x + \frac{c_1}{2}x^2 + \frac{c_2}{3}x^3 + \cdots + \frac{c_n}{n+1}x^{n+1}$. The null space of the differentiation transformation is the set of constant polynomials since $(Dc)(x) = 0$ for constants $c \in F$.

The range of the integration transformation is all polynomials with constant term equal to zero. Let $f(x) = c_1x + c_2x^2 + \cdots + x_nc^n$. Then $f(x) = (Tg)(x)$ where $g(x) = c_1 + 2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1}$. Clearly the integral transoformation of a polynomial has constant term equal to zero, so this is the entire range of the integration transformation. The null space of the integration transformation is the zero space $\{0\}$ since the (indefinite) integral of any other polynomial is non-zero.

Exercise 4: Is there a linear transformation T from $\mathbb{R}^3$ into $\mathbb{R}^2$ such that $T(1,-1,1) = (1,0)$ and $T(1,1,1) = (0,1)$?

Solution: Yes, there is such a linear transformation. Clearly $\alpha_1 = (1,-1,1)$ and $\alpha_2 = (1,1,1)$ are linearly independent. By Corollary 2, page 46, $\exists$ a third vector α_3 such that $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis for $\mathbb{R}^3$. By Theorem 1, page 69, there is a linear transformation that takes $\alpha_1, \alpha_2, \alpha_3$ to any three vectors we want. Therefore we can find a linear transformation that takes $\alpha_1 \mapsto (1,0)$, $\alpha_2 \mapsto (0,1)$ and $\alpha_3 \mapsto (0,0)$. (We could have used any vector instead of $(0,0)$.)

Exercise 5: If

$$\begin{aligned} \alpha_1 &= (1,-1), & \beta_1 &= (1,0) \\ \alpha_2 &= (2,-1), & \beta_2 &= (0,1) \\ \alpha_3 &= (-3,2), & \beta_3 &= (1,1) \end{aligned}$$

is there a linear transformation T from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that $T\alpha_i = \beta_i$ for $i = 1, 2$ and 3?

Solution: No there is no such transformation. If there was then since $\{\alpha_1, \alpha_2\}$ is a basis for $\mathbb{R}^2$ their images determine T completely. Now $\alpha_3 = -\alpha_1 - \alpha_2$, thus it must be that $T(\alpha_3) = T(-\alpha_1 - \alpha_2) = -T(\alpha_1) - T(\alpha_2) = -(1,0) - (0,1) = (-1,-1) \neq (1,1)$. Thus no such T can exist.

Exercise 6: Describe explicitly (as in Exercises 1 and 2) the linear transformation T from F^2 into F^2 such that $T\epsilon_1 = (a,b)$, $T\epsilon_2 = (c,d)$.

Solution: I'm not 100% sure I understand what they want here. Let A be the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then the range of T is the row-space of A which can have dimension 0, 1, or 2 depending on the row-rank. Explicitly it is all vectors of the form $x(a,b) + y(c,d) = (ax + cy, bx + dy)$ where x, y are arbitrary elements of F. The rank is the dimension of this row-space, which is 0 if $a = b = c = d = 0$ and if not all a, b, c, d are zero then by Exercise 1.6.8, page 27, the rank is 2 if $ad - bc \neq 0$ and equals 1 if $ad - bc = 0$.

Now let A be the matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Then the null space is the solution space of $AX = 0$. Thus the nullity is 2 if $a = b = c = d = 0$, and if not all a, b, c, d are zero then by Exercise 1.6.8, page 27 and Theorem 13, page 23, is 0 if $ad - bc \neq 0$ and is 1 if $ad - bc = 0$.

Exercise 7: Let F be a subfield of the complex numbers and let T be the function from F^3 into F^3 defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3).$$

(a) Verify that T is a linear transformation.

(b) If (a, b, c) is a vector in F^3, what are the conditions on a, b, and c that the vector be in the range of T? What is the rank of T?

(c) What are the conditions on a, b, and c that (a, b, c) be in the null space of T? What is the nullity of T?

Solution: (a) Let

$$P = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ -1 & -2 & 2 \end{bmatrix}.$$

Then T can be represented by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto P \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

By Example 4, page 68, this is a linear transformation, where we've identified F^3 with $F^{3\times 1}$ and taken Q in Example 4 to be the identity matrix.

(b) The range of T is the column space of P, or equivalently the row space of

$$P^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & -2 \\ 2 & 0 & 2 \end{bmatrix}.$$

We row reduce the matrix as follows

$$\rightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -3 \\ 0 & -4 & 4 \end{bmatrix}.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let $\rho_1 = (1, 0, 1)$ and $\rho_2 = (0, 1, -1)$. Then elements of the row space are elements of the form $b_1\rho_1 + b_2\rho_2 = (b_1, b_2, b_1 - b_2)$. Thus the rank of T is two and (a, b, c) is in the range of T as long as $c = a - b$.

Alternatively, we can row reduce the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & a \\ 2 & 1 & 0 & b \\ -1 & -2 & 2 & c \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 2 & a \\ 0 & 3 & -4 & b-2a \\ 0 & -3 & 4 & a+c \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 2 & a \\ 0 & 3 & -4 & b-2a \\ 0 & 0 & 0 & -a+b+c \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 2 & a \\ 0 & 1 & -4/3 & (b-2a)/4 \\ 0 & 0 & 0 & -a+b+c \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2/3 & (b+2a)/4 \\ 0 & 1 & -4/3 & (b-2a)/4 \\ 0 & 0 & 0 & -a+b+c \end{array}\right]$$

from which we arrive at the condition $-a+b+c=0$ or equivalently $c=a-b$.

(c) We must find all $X = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ such that $PX = 0$ where P is the matrix from part (a). We row reduce the matrix

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ -1 & -2 & 2 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -4 \\ 0 & -3 & 4 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -4/3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 2/3 \\ 0 & 1 & -4/3 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore

$$\begin{cases} a + \frac{2}{3}c = 0 \\ b - \frac{4}{3}c = 0 \end{cases}$$

So elements of the null space of T are of the form $(-\frac{2}{3}c, \frac{4}{3}c, c)$ for arbitrary $c \in F$ and the dimension of the null space (the nullity) equals one.

Exercise 8: Describe explicitly a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^3$ which has as its range the subspace spanned by $(1,0,-1)$ and $(1,2,2)$.

Solution: By Theorem 1, page 69, (and its proof) there is a linear transformation T from $\mathbb{R}^3$ to $\mathbb{R}^3$ such that $T(1,0,0) = (1,0,-1)$, $T(0,1,0) = (1,0,-1)$ and $T(0,0,1) = (1,2,2)$ and the range of T is exactly the subspace generated by

$$\{T(1,0,0), T(0,1,0), T(0,0,1)\} = \{(1,0,-1), (1,2,2)\}.$$

Exercise 9: Let V be the vector space of all $n \times n$ matrices over the field F, and let B be a fixed $n \times n$ matrix. If

$$T(A) = AB - BA$$

verify that T is a linear transformation from V into V.

Solution: $T(cA_1 + A_2) = (cA_1 + A_2)B - B(cA_1 + A_2) = cA_1B + A_2B - cBA_1 - BA_2$

$$= c(A_1B - BA_1) + (A_2B - BA_2) = cT(A_1) + T(A_2).$$

Exercise 10: Let V be the set of all complex numbers regarded as a vector space over the field of real numbers (usual operations). Find a function from V into V which is a linear transformation on the above vector space, but which is not a linear transformation on $\mathbb{C}^1$, i.e., which is not complex linear.

Solution: Let $T : V \to V$ be given by $a + bi \mapsto a$. Let $z = a + bi$ and $w = a' + b'i$ and $c \in \mathbb{R}$. Then $T(cz + w) = T((ca + a') + (cb + b')i) = ca + a' = cT(a + bi) + T(a' + b'i) = aT(z) + T(w)$. Thus T is real linear. However, if T were complex linear then we must have $0 = T(i) = T(i \cdot 1) = i \cdot T(1) = i \cdot 1 = i$. But $0 \neq i$ so this is a contradiction. Thus T is not complex linear.

Exercise 11: Let V be the space of $n \times 1$ matrices over F and let W be the space of $m \times 1$ matrices over F. Let A be a fixed $m \times n$ matrix over F and let T be the linear transformation from V into W defined by $T(X) = AX$. Prove that T is the zero transformation if and only if A is the zero matrix.

Solution: If A is the zero matrix then clearly T is the zero transformation. Conversely, suppose A is not the zero matrix, suppose the k-th column A_k has a non-zero entry. Then $T(\epsilon_k) = A_k \neq 0$.

Exercise 12: Let V be an n-dimensional vector space over the field F and let T be a linear transformation from V into V such that the range and null space of T are identical. Prove that n is even. (Can you give an example of such a linear transformation T?)

Solution: From Theorem 2, page 71, we know $\text{rank}(T) + \text{nullity}(T) = \dim V$. In this case we are assuming both terms on the left hand side are equal, say equal to m. Thus $m + m = n$ or equivalently $n = 2m$ which implies n is even.

The simplest example is $V = \{0\}$ the zero space. Then trivially the range and null space are equal. To give a less trivial example assume $V = \mathbb{R}^2$ and define T by $T(1, 0) = (0, 0)$ and $T(0, 1) = (1, 0)$. We can do this by Theorem 1, page 69 because $\{(1, 0), (0, 1)\}$ is a basis for $\mathbb{R}^2$. Then clearly the range and null space are both equal to the subspace of $\mathbb{R}^2$ generated by $(1, 0)$.

Exercise 13: Let V be a vector space and T a linear transformation from V into V. Prove that the following two statements about T are equivalent.

(a) The intersection of the range of T and the null space of T is the zero subspace of V.

(b) If $T(T\alpha) = 0$, then $T\alpha = 0$.

Solution: (a) $\Rightarrow$ (b): Statement (a) says that nothing in the range gets mapped to zero except for 0. In other words if x is in the range of T then $Tx = 0 \Rightarrow x = 0$. Now $T\alpha$ is in the range of T, thus $T(T\alpha) = 0 \Rightarrow T\alpha = 0$.

(b) $\Rightarrow$ (a): Suppose x is in both the range and null space of T. Since x is in the range, $x = T\alpha$ for some α. But then x in the null space of T implies $T(x) = 0$ which implies $T(T\alpha) = 0$. Thus statement (b) implies $T\alpha = 0$ or equivalently $x = 0$. Thus the only thing in both the range and null space of T is the zero vector 0.

Section 3.2: The Algebra of Linear Transformations

Page 76: Typo in line 1: It says $A_{ij}, \dots, A_{mj}$, it should say $A_{1j}, \dots, A_{mj}$.

Exercise 1: Let T and U be the linear operators on $\mathbb{R}^2$ defined by

$$T(x_1, x_2) = (x_2, x_1) \quad \text{and} \quad U(x_1, x_2) = (x_1, 0).$$

(a) How would you describe T and U geometrically?

(b) Give rules like the ones defining T and U for each of the transformations $(U + T)$, UT, TU, T^2, U^2.

Solution: (a) Geometrically, in the $x-y$ plane, T is the reflection about the diagonal $x = y$ and U is a projection onto the x-axis.

(b)

- $(U + T)(x_1, x_2) = (x_2, x_1) + (x_1, 0) = (x_1 + x_2, x_1)$.
- $(UT)(x_1, x_2) = U(x_2, x_1) = (x_2, 0)$.
- $(TU)(x_1, x_2) = T(x_1, 0) = (0, x_1)$.
- $T^2(x_1, x_2) = T(x_2, x_1) = (x_1, x_2)$, the identity function.
- $U^2(x_1, x_2) = U(x_1, 0) = (x_1, 0)$. So $U^2 = U$.

Exercise 2: Let T be the (unique) linear operator on $\mathbb{C}^3$ for which

$$T\epsilon_1 = (1, 0, i), \quad T\epsilon_2 = (0, 1, 1), \quad T\epsilon_3 = (i, 1, 0).$$

Is T invertible?

Solution: By Theorem 9 part (v), top of page 82, T is invertible if $\{T\epsilon_1, T\epsilon_2, T\epsilon_3\}$ is a basis of $\mathbb{C}^3$. Since $\mathbb{C}^3$ has dimension three, it suffices (by Corollary 1 page 46) to show $T\epsilon_1, T\epsilon_2, T\epsilon_3$ are linearly independent. To do this we row reduce the matrix

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

to row-reduced echelon form. If it reduces to the identity then its rows are independent, otherwise they are dependent. Row reduction follows:

$$\begin{bmatrix} 1 & 0 & i \\ 0 & 1 & 1 \\ i & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & i \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & i \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

This is in row-reduced echelon form not equal to the identity. Thus T is not invertible.

Exercise 3: Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3).$$

Is T invertible? If so, find a rule for T^{-1} like the one which defines T.

Solution: The matrix representation of the transformation is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 3 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

where we've identified $\mathbb{R}^3$ with $\mathbb{R}^{3\times 1}$. T is invertible if the matrix of the transformation is invertible. To determine this we row-reduce the matrix - we row-reduce the augmented matrix to determine the inverse for the second part of the Exercise.

$$\left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 & -3 & 0 \\ 0 & 3 & 1 & 0 & -2 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1/3 & -1 & 0 \\ 0 & 3 & 1 & 0 & -2 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & 0 & 0 \\ 0 & 1 & 0 & 1/3 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]$$

Since the left side transformed into the identity, T is invertible. The inverse transformation is given by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1/3 & 0 & 0 \\ 1/3 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

So

$$T^{-1}(x_1, x_2, x_3) = (x_1/3,\ \ x_1/3 - x_2,\ \ -x_1 + x_2 + x_3).$$

Exercise 4: For the linear operator T of Exercise 3, prove that

$$(T^2 - I)(T - 3I) = 0.$$

Solution: Working with the matrix representation of T we must show

$$(A^2 - I)(A - 3I) = 0$$

where

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & 1 \end{bmatrix}.$$

Calculating:

$$A^2 = \begin{bmatrix} 3 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 9 & 0 & 0 \\ 2 & 1 & 0 \\ 9 & 0 & 1 \end{bmatrix}$$

Thus

$$A^2 - I = \begin{bmatrix} 8 & 0 & 0 \\ 2 & 0 & 0 \\ 9 & 0 & 0 \end{bmatrix}.$$

Also

$$A - 3I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -4 & 0 \\ 2 & 1 & -2 \end{bmatrix}$$

Thus

$$(A^2 - I)(A - 3I) = \begin{bmatrix} 8 & 0 & 0 \\ 2 & 0 & 0 \\ 9 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 \\ 1 & -4 & 0 \\ 2 & 1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Exercise 5: Let $C^{2\times 2}$ be the complex vector space of 2×2 matrices with complex entries. Let

$$B = \begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix}$$

and let T be the linear operator on $\mathbb{C}^{2\times 2}$ defined by $T(A) = BA$. What is the rank of T? Can you describe T^2?

Solution: An (ordered) basis for $\mathbb{C}^{2\times 2}$ is given by

$$A_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$A_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

If we identify $\mathbb{C}^{2\times 2}$ with $\mathbb{C}^4$ by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto (a, b, c, d)$$

then since

$$A_{11} \mapsto A_{11} - 4A_{21}$$

$$A_{21} \mapsto -A_{11} + 4A_{21}$$

$$A_{12} \mapsto A_{12} - 4A_{22}$$

$$A_{22} \mapsto -A_{12} + 4A_{22}$$

the matrix of the transformation is given by

$$\begin{bmatrix} 1 & -4 & 0 & 0 \\ -1 & 4 & 0 & 0 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & -1 & 4 \end{bmatrix}.$$

To find the rank of T we row-reduce this matrix:

$$\to \begin{bmatrix} 1 & -4 & 0 & 0 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It has rank two so the rank, so the rank of T is 2.

$T^2(A) = T(T(A)) = T(BA) = B(BA) = B^2A$. Thus T^2 is given by multiplication by a matrix just as T is, but multiplication with B^2 instead of B. Explicitly

$$B^2 = \begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & -5 \\ -20 & 20 \end{bmatrix}.$$

Exercise 6: Let T be a linear transformation from $\mathbb{R}^3$ into $\mathbb{R}^2$, and let U be a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$. Prove that the transformation UT is not invertible. Generalize the theorem.

Solution: Let $\{\alpha_1, \alpha_2, \alpha_3\}$ be a basis for $\mathbb{R}^3$. Then $T(\alpha_1), T(\alpha_2), T(\alpha_3)$ must be linearly dependent in $\mathbb{R}^2$, because $\mathbb{R}^2$ has dimension 2. So suppose $b_1T(\alpha_1) + b_2T(\alpha_2) + b_3T(\alpha_3) = 0$ and not all b_1, b_2, b_3 are zero. Then $b_1\alpha_1 + b_2\alpha_2 + b_3\alpha_3 \neq 0$ and

$$UT(b_1\alpha_1 + b_2\alpha_2 + b_3\alpha_3)$$

$$= U(T(b_1\alpha_1 + b_2\alpha_2 + b_3\alpha_3))$$

$$= U(b_1T(\alpha_1) + b_2T(\alpha_2) + b_3T(\alpha_3)$$

$$= U(0) = 0.$$

Thus (by the definition at the bottom of page 79) UT is *not* non-singular and thus by Theorem 9, page 81, UT is not invertible.

The obvious generalization is that if $n > m$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^m \to \mathbb{R}^n$ are linear transformations, then UT is not invertible. The proof is an immediate generalization the proof of the special case above, just replace α_3 with $\dots, \alpha_n$.

Exercise 7: Find two linear operators T and U on $\mathbb{R}^2$ such that $TU = 0$ but $UT \neq 0$.

Solution: Identify $\mathbb{R}^2$ with $\mathbb{R}^{2\times 1}$ and let T and U be given by the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

More precisely, for

$$X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

let T be given by $X \mapsto AX$ and let U be given by $X \mapsto BX$. Thus TU is given by $X \mapsto ABX$ and UT is given by $X \mapsto BAX$. But $BA = 0$ and $AB \neq 0$ so we have the desired example.

Exercise 8: Let V be a vector space over the field F and T a linear operator on V. If $T^2 = 0$, what can you say about the relation of the range of T to the null space of T? Give an example of a linear operator T on $\mathbb{R}^2$ such that $T^2 = 0$ but $T \neq 0$.

Solution: If $T^2 = 0$ then the range of T must be contained in the null space of T since if y is in the range of T then $y = Tx$ for some x so $Ty = T(Tx) = T^2x = 0$. Thus y is in the null space of T.

To give an example of an operator where $T^2 = 0$ but $T \neq 0$, let $V = \mathbb{R}^{2\times 1}$ and let T be given by the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Specifically, for

$$X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

let T be given by $X \mapsto AX$. Since $A \neq 0$, $T \neq 0$. Now T^2 is given by $X \mapsto A^2X$, but $A^2 = 0$. Thus $T^2 = 0$.

Exercise 9: Let T be a linear operator on the finite-dimensional space V. Suppose there is a linear operator U on V such that $TU = I$. Prove that T is invertible and $U = T^{-1}$. Give an example which shows that this is false when V is not finite-dimensional. (*Hint:* Let $T = D$, be the differentiation operator on the space of polynomial functions.)

Solution: By the comments in the Appendix on functions, at the bottom of page 389, we see that simply because $TU = I$ as functions, then necessarily T is onto and U is one-to-one. It then follows immediately from Theorem 9, page 81, that T is invertible. Now $TT^{-1} = I = TU$ and multiplying on the left by T^{-1} we get $T^{-1}TT^{-1} = T^{-1}TU$ which implies $(I)T^{-1} = (I)U$ and thus $U = T^{-1}$.

Let V be the space of polynomial functions in one variable over $\mathbb{R}$. Let D be the differentiation operator and let T be the operator "multiplication by x" (exactly as in Example 11, page 80). As shown in Example 11, $UT = I$ while $TU \neq I$. Thus this example fulfills the requirement.

Exercise 10: Let A be an $m \times n$ matrix with entries in F and let T be the linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$ defined by $TX = AX$. Show that if $m < n$ it may happen that T is onto without being non-singular. Similarly, show that if $m > n$ we may have T non-singular but not onto.

Solution: Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ be a basis for $F^{n\times 1}$ and let $\mathcal{B}' = \{\beta_1, \dots, \beta_m\}$ be a basis for $F^{m\times 1}$. We can define a linear transformation from $F^{n\times 1}$ to $F^{m\times 1}$ uniquely by specifying where each member of $\mathcal{B}$ goes in $F^{m\times 1}$. If $m < n$ then we can define a linear transformation that maps at least one member of $\mathcal{B}$ to each member of $\mathcal{B}'$ and maps at least two members of $\mathcal{B}$ to the same member of $\mathcal{B}'$. Any linear transformation so defined must necessarily be onto without being one-to-one. Similarly, if $m > n$ then we can map each member of $\mathcal{B}$ to a unique member of $\mathcal{B}'$ with at least one member of $\mathcal{B}'$ not mapped to by any member of $\mathcal{B}$. Any such transformation so defined will necessarily be one-to-one but not onto.

Exercise 11: Let V be a finite-dimensional vector space and let T be a linear operator on V. Suppose that $\text{rank}(T^2) = \text{rank}(T)$. Prove that the range and null space of T are disjoint, i.e., have only the zero vector in common.

Solution: Let $\{\alpha_1, \dots, \alpha_n\}$ be a basis for V. Then the rank of T is the number of linearly independent vectors in the set $\{T\alpha_1, \dots, T\alpha_n\}$. Suppose the rank of T equals k and suppose WLOG that $\{T\alpha_1, \dots, T\alpha_k\}$ is a linearly independent set (it might be that $k = 1$, pardon the notation). Then $\{T\alpha_1, \dots, T\alpha_k\}$ give a basis for the range of T. It follows that $\{T^2\alpha_1, \dots, T^2\alpha_k\}$ span the range of T^2 and since the dimension of the range of T^2 is also equal to k, $\{T^2\alpha_1, \dots, T^2\alpha_k\}$ must be a basis for the range of T^2. Now suppose v is in the range of T. Then $v = c_1T\alpha_1 + \cdots + c_kT\alpha_k$. Suppose v is also in the null space of T. Then $0 = T(v) = T(c_1T\alpha_1 + \cdots + c_kT\alpha_k) = c_1T^2\alpha_1 + \cdots + c_kT^2\alpha_k$. But $\{T^2\alpha_1, \dots, T^2\alpha_k\}$ is a basis, so $T^2\alpha_1, \dots, T^2\alpha_k$ are linearly independent, thus it must be that $c_1 = \cdots = c_k = 0$, which implies $v = 0$. Thus we have shown that if v is in both the range of T and the null space of T then $v = 0$, as required.

Exercise 12: Let p, m, and n be positive integers and F a field. Let V be the space of $m \times n$ matrices over F and W the space of $p \times n$ matrices over F. Let B be a fixed $p \times m$ matrix and let T be the linear transformation from V into W defined by $T(A) = BA$. Prove that T is invertible if and only if $p = m$ and B is an invertible $m \times m$ matrix.

Solution: We showed in Exercise 2.3.12, page 49, that the dimension of V is mn and the dimension of W is pn. By Theorem 9 page (iv) we know that an invertible linear transformation must take a basis to a basis. Thus if there's an invertible linear transformation between V and W it must be that both spaces have the same dimension. Thus if T is inverible then $pn = mn$ which implies $p = m$. The matrix B is then invertible because the assignment $B \mapsto BX$ is one-to-one (Theorem 9 (ii), page 81) and non-invertible matrices have non-trivial solutions to $BX = 0$ (Theorem 13, page 23). Conversely, if $p = n$ and B is invertible, then we can define the inverse transformation T^{-1} by $T^{-1}(A) = B^{-1}A$ and it follows that T is invertible.

Section 3.3: Isomorphism

Exercise 1: Let V be the set of complex numbers and let F be the field of real numbers. With the usual operations, V is a vector space over F. Describe explicitly an isomorphism of this space onto $\mathbb{R}^2$.

Solution: The natural isomorphism from V to $\mathbb{R}^2$ is given by $a + bi \mapsto (a, b)$. Since i acts like a placeholder for addition in $\mathbb{C}$, $(a + bi) + (c + di) = (a + c) + (b + d)i \mapsto (a + c, b + d) = (a, b) + (c, d)$. And $c(a + bi) = ca + cbi \mapsto (ca, cb) = c(a, b)$. Thus this is a linear transformation. The inverse is clearly $(a, b) \mapsto a + bi$. Thus the two spaces are isomorphic as vector spaces over $\mathbb{R}$.

Exercise 2: Let V be a vector space over the field of complex numbers, and suppose there is an isomorphism T of V into $\mathbb{C}^3$. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be vectors in V such that

$$T\alpha_1 = (1, 0, i), \quad T\alpha_2 = (-2, 1 + i, 0),$$

$$T\alpha_3 = (-1, 1, 1), \quad T\alpha_4 = (\sqrt{2}, i, 3).$$

(a) Is α_1 in the subspace spanned by α_2 and α_3?

(b) Let W_1 be the subspace spanned by α_1 and α_2, and let W_2 be the subspace spanned by α_3 and α_4. What is the intersection of W_1 and W_2?

(c) Find a basis for the subspace of V spanned by the four vectors α_j.

Solution: (a) Since T is an isomorphism, it suffices to determine whether $T\alpha_1$ is contained in the subspace spanned by $T\alpha_2$ and $T\alpha_3$. In other words we need to determine if there is a solution to

$$\left[\begin{array}{cc} -2 & -1 \\ 1+i & 1 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} 1 \\ 0 \\ i \end{array}\right].$$

To do this we row-reduce the augmented matrix

$$\left[\begin{array}{cc|c} -2 & -1 & 1 \\ 1+i & 1 & 0 \\ 0 & 1 & i \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1/2 & -1/2 \\ 1+i & 1 & 0 \\ 0 & 1 & i \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1/2 & -1/2 \\ 0 & 1 & i \\ 1+i & 1 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|c} 1 & 1/2 & -1/2 \\ 0 & 1 & i \\ 0 & \frac{1-i}{2} & \frac{1+i}{2} \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & \frac{-1-i}{2} \\ 0 & 1 & i \\ 0 & 0 & 0 \end{array}\right]$$

The zero row on the left of the dividing line has zero also on the right. This means the system has a solution. Therefore we can conclude that α_1 is in the subspace generated by α_2 and α_3.

(b) Since $T\alpha_1$ and $T\alpha_2$ are linearly independent, and $T\alpha_3$ and $T\alpha_4$ are linearly independent, $\dim(W_1) = \dim(W_2) = 2$. We row-reduce the matrix whose columns are the $T\alpha_i$:

$$\left[\begin{array}{cccc} 1 & -2 & -1 & \sqrt{2} \\ 0 & 1+i & 1 & i \\ i & 0 & 1 & 3 \end{array}\right]$$

which yields

$$\left[\begin{array}{cccc} 1 & 0 & -i & 0 \\ 0 & 1 & \frac{1-i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right],$$

from which we deduce that $T\alpha_1, T\alpha_2, T\alpha_3, T\alpha_4$ generate a space of dimension three, thus $\dim(W_1+W_2) = 3$. Since $\dim(W_1) = \dim(W_2) = 2$ it follows from Theorem 6, page 46 that $\dim(W_1 \cap W_2) = 1$. Now $AX = 0 \Leftrightarrow RX = 0$ where R is the row reduced echelon form of A. This follows from the fact that $R = PA$; multiply both sides of $AX = 0$ on the left by P. Solving for X in $RX = 0$ gives the general solution is of the form $(ic, \frac{i-1}{2}c, c, 0)$. Letting $c = 2$ gives

$$2iT\alpha_1 + (i-1)T\alpha_2 + 2T\alpha_3 = 0$$

which implies $T\alpha_3 = -iT\alpha_1 + \frac{1-i}{2}T\alpha_2$ which implies $T\alpha_3 \in TW_1$. Thus $\alpha_3 \in W_1$. Thus $\alpha_3 \in W_1 \cap W_2$. Since $\dim(W_1 \cap W_2) = 1$ it follows that $W_1 \cap W_2 = \mathbb{C}\alpha_3$.

(c) We have determined in part (b) that the $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ span a space of dimension three, and that α_3 is in the space generated by α_1 and α_2. Thus $\{\alpha_1, \alpha_2, \alpha_4\}$ give a basis for the subspace spanned by $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, which in fact is all of $\mathbb{C}^3$.

Exercise 3: Let W be the set of all 2×2 complex Hermitian matrices, that is, the set of 2×2 complex matrices A such that $A_{ij} = \overline{A_{ji}}$ (the bar denoting complex conjugation). As we pointed out in Example 6 of Chapter 2, W is a vector space over the field of *real* numbers, under the usual operations. Verify that

$$(x, y, z, t) \rightarrow \begin{bmatrix} t+x & y+iz \\ y-iz & t-x \end{bmatrix}$$

is an isomorphism of $\mathbb{R}^4$ onto W.

Solution: The function is linear since the four components are all linear combinations of the components of the domain (x, y, z, t). Identify $\mathbb{C}^{2\times2}$ with $\mathbb{C}^4$ by $A \mapsto (A_{11}, A_{12}, A_{21}, A_{22})$. Then the matrix of the transformation is given by

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

As usual, the transformation is an isomorphism if the matrix is invertible. We row-reduce to veryify the matrix is invertible. We will row-reduce the augmented matrix in order to find the inverse explicitly:

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & i & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -i & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

This reduces to

$$\rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & -1/2 \\ 0 & 1 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & -i/2 & i/2 & 0 \\ 0 & 0 & 0 & 1 & 1/2 & 0 & 0 & 1/2 \end{array}\right].$$

Thus the inverse transformation is

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \mapsto \left(\frac{x-w}{2}, \ \frac{y+z}{2}, \ \frac{i(z-y)}{2}, \ \frac{x+w}{2} \right).$$

Exercise 4: Show that $F^{m\times n}$ is isomorphic to F^{mn}.

Solution: Define the bijection σ from $\{(a, b) \mid a, b \in \mathbb{N}, 1 \leq a \leq m, 1 \leq b \leq n\}$ to $\{1, 2, \ldots, mn\}$ by $(a, b) \mapsto (a-1)n+b$. Define the function G from $F^{m\times n}$ to F^{mn} as follows. Let $A \in F^{m\times n}$. Then map A to the mn-tuple that has A_{ij} in the $\sigma(i, j)$ position. In other words $A \mapsto (A_{11}, A_{12}, A_{13}, \ldots, A_{1n}, A_{21}, A_{22}, A_{23}, \ldots, A_{2n}, \ldots\ldots, A_{nn})$. Since addition in $F^{m\times n}$ and in F^{mn} is performed compenent-wise, $G(A + B) = G(A) + G(B)$. Similarly since scalar multiplication factors out of vectors component-wise in the same way in $F^{m\times n}$ as in F^{mn}, we also have $G(cA) = cG(A)$. Thus G is a linear function. G is clearly one-to-one (as well as clearly onto), and both $F^{m\times n}$ and F^{mn} have dimension mn (by Example 17, page 45 and Exercise 2.3.12, page 49), thus (by Theorem 9, page 81) it follows that G has an inverse and therefore is an isomorphism.

Exercise 5: Let V be the set of complex numbers regarded as a vector space over the field of real numbers (Exercise 1). We define a function T from V into the space of 2×2 real matrices, as follows. If $z = x + iy$ with x and y real numbers, then

$$T(z) = \begin{bmatrix} x+7y & 5y \\ -10y & x-7y \end{bmatrix}.$$

(a) Verify that T is a one-one (real) linear transformation of V into the space of 2×2 matrices.

(b) Verify that $T(z_1 z_2) = T(z_1)T(z_2)$.

(c) How would you describe the range of T?

Solution:

(a) The four coordinates of $T(z)$ are written as linear combinations of the coordinates of z (as a vector over $\mathbb{R}$). Thus T is clearly a linear transformation. To see that T is one-to-one, let $z = x + yi$ and $w = a + bi$ and suppose $T(z) = T(w)$. Then considering the top right entry of the matrix we see that $5y = 5b$ which implies $b = y$. It now follows from the top left entry of the matrix that $x = a$. Thus $T(z) = T(w) \Rightarrow z = w$, thus T is one-to-one.

(b) Let $z_1 = x + yi$ and $z_2 = a + bi$. Then

$$T(z_1z_2) = T((ax - by) + (ay + bx)i) = \begin{bmatrix} (ax - by) + 7(ay + bx) & 5(ay + bx) \\ -10(ay + bx) & (ax - by) - 7(ay + bx) \end{bmatrix}.$$

On the other hand,

$$T(z_1)T(z_2) = \begin{bmatrix} x + 7y & 5y \\ -10y & x - 7y \end{bmatrix}\begin{bmatrix} a + 7b & 5b \\ -10b & a - 7b \end{bmatrix}$$

$$= \begin{bmatrix} (ax - by) + 7(ay + bx) & 5(ay + bx) \\ -10(ay + bx) & (ax - by) - 7(ay + bx) \end{bmatrix}.$$

Thus $T(z_1z_2) = T(z_1)T(z_2)$.

(c) The range of T has (real) dimension equal to two by part (a), and so the range of T is isomorphic to $\mathbb{C}$ as *real* vector spaces. But both spaces also have a natural multiplication and in part (b) we showed that T respects the multiplication. Thus the range of T is isomorphic to $\mathbb{C}$ as fields and we have essentially found an isomorphic copy of the field $\mathbb{C}$ in the algebra of 2×2 real matrices.

Exercise 6: Let V and W be finite-dimensional vector spaces over the field F. Prove that V and W are isomorphic if and only if $\dim(V) = \dim(W)$.

Solution: Suppose $\dim(V) = \dim(W) = n$. By Theorem 10, page 84, both V and W are isomorphic to F^n, and consequently, since isomorphism is an equivalence relation, V and W are isomorphic to each other. Conversely, suppose T is an isomorphism from V to W. Suppose $\dim(W) = n$. Then by Theorem 10 again, there is an isomorphism $S : W \to F^n$. Thus ST is an isomorphism from V to F^n implying also $\dim(V) = n$.

Exercise 7: Let V and W be vector spaces over the field F and let U be an isomorphism of V onto W. Prove that $T \to UTU^{-1}$ is an isomorphism of $L(V, V)$ onto $L(W, W)$.

Solution: $L(V, V)$ is defined on page 75 as the vector space of linear transformations from V to V, and likewise $L(W, W)$ is the vector space of linear transformations from W to W.

Call the function f. We know $f(T)$ is linear since it is a composition of three linear tranformations UTU^{-1}. Thus indeed f is a function from $L(V, V)$ to $L(W, W)$. Now $f(aT + T') = U(aT + T')U^{-1} = (aUT + UT')U^{-1} = aUTU^{-1} + UT'U^{-1} = af(T) + f(T')$. Thus f is linear. We just must show f has an inverse. Let g be the function from $L(W, W)$ to $L(V, V)$ given by $g(T) = U^{-1}TU$. Then $gf(T) = U^{-1}(UTU^{-1})U = T$. Similarly $fg = I$. Thus f and g are inverses. Thus f is an isomorphism.

Section 3.4: Representation of Transformations by Matrices

Page 90: Typo. Four lines from the bottom it says "Example 12" where they probably meant Example 10 (page 78).

Page 91: Just before (3-8) it says "By definition". I think it's more than just by definition, see bottom of page 88.

Exercise 1: Let T be the linear operator on $\mathbb{C}^2$ defined by $T(x_1, x_2) = (x_1, 0)$. Let $\mathcal{B}$ be the standard ordered basis for $\mathbb{C}^2$ and let $\mathcal{B}' = \{\alpha_1, \alpha_2\}$ be the ordered basis defined by $\alpha_1 = (1, i)$, $\alpha_2 = (-i, 2)$.

(a) What is the matrix of T relative to the pair $\mathcal{B}, \mathcal{B}'$?

(b) What is the matrix of T relative to the pair $\mathcal{B}', \mathcal{B}$?

(c) What is the matrix of T in the ordered basis $\mathcal{B}'$?

(d) What is the matrix of T in the ordered basis $\{\alpha_2, \alpha_1\}$?

Solution: (a) According to the comments at the bottom of page 87, the i-th column of the matrix is given by $[T\epsilon_i]_{\mathcal{B}'}$, where $\epsilon_1 = (1, 0)$ and $\epsilon_2 = (0, 1)$, the standard basis vectors of $\mathbb{C}^2$. Now $T\epsilon_1 = (1, 0)$ and $T\epsilon_2 = (0, 0)$. To write these in terms of α_1 and α_2 we use the approach of row-reducing the augmented matrix

$$\left[\begin{array}{cc|cc} 1 & -i & 1 & 0 \\ i & 2 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & -i & 1 & 0 \\ 0 & 1 & -i & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 2 & 0 \\ 0 & 1 & -i & 0 \end{array}\right].$$

Thus $T\epsilon_1 = 2\alpha_1 - i\alpha_2$ and $T\epsilon_2 = 0 \cdot \alpha_1 + 0 \cdot \alpha_2$ and the matrix of T relative to $\mathcal{B}, \mathcal{B}'$ is

$$\left[\begin{array}{cc} 2 & 0 \\ -i & 0 \end{array}\right].$$

(b) In this case we have to write $T\alpha_1$ and $T\alpha_2$ as linear combinations of ϵ_1, ϵ_2.

$$T\alpha_1 = (1, 0) = 1 \cdot \epsilon_1 + 0 \cdot \epsilon_2$$

$$T\alpha_2 = (-i, 0) = -i \cdot \epsilon_1 + 0 \cdot \epsilon_2.$$

Thus the matrix of T relative to $\mathcal{B}', \mathcal{B}$ is

$$\left[\begin{array}{cc} 1 & -i \\ 0 & 0 \end{array}\right].$$

(c) In this case we need to write $T\alpha_1$ and $T\alpha_2$ as linear combinations of α_1 and α_2. $T\alpha_1 = (1, 0)$, $T\alpha_2 = (-i, 0)$. We row-reduce the augmented matrix:

$$\left[\begin{array}{cc|cc} 1 & -i & 1 & -i \\ i & 2 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & -i & 1 & -i \\ 0 & 1 & -i & -1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 2 & -2i \\ 0 & 1 & -i & -1 \end{array}\right].$$

Thus the matrix of T in the ordered basis $\mathcal{B}'$ is

$$\left[\begin{array}{cc} 2 & -2i \\ -i & -1 \end{array}\right].$$

(d) In this case we need to write $T\alpha_2$ and $T\alpha_1$ as linear combinations of α_2 and α_1. In this case the matrix we need to row-reduce is just the same as in (c) but with columns switched:

$$\left[\begin{array}{cc|cc} -i & 1 & -i & 1 \\ 2 & i & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & i & 1 & i \\ 2 & i & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & i & 1 & i \\ 0 & -i & -2 & -2i \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & i & 1 & i \\ 0 & 1 & -2i & 2 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -1 & -i \\ 0 & 1 & -2i & 2 \end{array}\right]$$

Thus the matrix of T in the ordered basis $\{\alpha_2, \alpha_1\}$ is

$$\left[\begin{array}{cc} -1 & -i \\ -2i & 2 \end{array}\right].$$

Exercise 2: Let T be the linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1).$$

(a) If $\mathcal{B}$ is the standard ordered basisfor $\mathbb{R}^3$ and $\mathcal{B}'$ is the standard ordered basis for $\mathbb{R}^2$, what is the matrix of T relative to the pair $\mathcal{B}, \mathcal{B}'$?

(b) If $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ and $\mathcal{B}' = \{\beta_1, \beta_2\}$, where

$$\alpha_1 = (1, 0, -1), \quad \alpha_2 = (1, 1, 1), \quad \alpha_3 = (1, 0, 0), \quad \beta_1 = (0, 1), \quad \beta_2 = (1, 0)$$

what is the matrix of T relative to the pair $\mathcal{B}, \mathcal{B}'$?

Solution: With respect to the standard bases, the matrix is simply

$$\left[\begin{array}{ccc} 1 & 1 & 0 \\ -1 & 0 & 2 \end{array}\right].$$

(b) We must write $T\alpha_1, T\alpha_2, T\alpha_3$ in terms of β_1, β_2.

$$T\alpha_1 = (1, -3)$$

$$T\alpha_2 = (2, 1)$$

$$T\alpha_3 = (1, 0).$$

We row-reduce the augmented matrix

$$\left[\begin{array}{cc|ccc} 0 & 1 & 1 & 2 & 1 \\ 1 & 0 & -3 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cc|ccc} 1 & 0 & -3 & 1 & 0 \\ 0 & 1 & 1 & 2 & 1 \end{array}\right].$$

Thus the matrix of T with respect to $\mathcal{B}, \mathcal{B}'$ is

$$\left[\begin{array}{ccc} -3 & 1 & 0 \\ 1 & 2 & 1 \end{array}\right].$$

Exercise 3: Let T be a linear operator on F^n, let A be the matrix of T in the standard ordered basis for F^n, and let W be the subspace of F^n spanned by the column vectors of A. What does W have to do with T?

Solution: Since $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of F^n, we know $\{T\epsilon_1, \ldots, T\epsilon_n\}$ generate the range of T. But $T\epsilon_i$ equals the i-th column vector of A. Thus the column vectors of A generate the range of T (where we identify F^n with $F^{n\times 1}$). We can also conclude that a subset of the columns of A give a basis for the range of T.

Exercise 4: Let V be a two-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis for V. If T is a linear operator on V and

$$[T]_{\mathcal{B}} = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$$

prove that $T^2 - (a + d)T + (ad - bc)I = 0$.

Solution: The coordinate matrix of $T^2 - (a + d)T + (ad - bc)I$ with respect to $\mathcal{B}$ is

$$[T^2 - (a + d)T + (ad - bc)I]_{\mathcal{B}} = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]^2 - (a + d)\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] + (ad - bc)\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$$

Expanding gives

$$= \left[\begin{array}{cc} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{array}\right] - \left[\begin{array}{cc} a^2 + ad & ab + bd \\ ac + cd & ad + d^2 \end{array}\right] + \left[\begin{array}{cc} ad - bc & 0 \\ 0 & ad - bc \end{array}\right]$$

$$= \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right].$$

Thus $T^2 - (a+d)T + (ad-bc)I$ is represented by the zero matrix with respect to $\mathcal{B}$. Thus $T^2 - (a+d)T + (ad-bc)I = 0$.

Exercise 5: Let T be the linear operator on $\mathbb{R}^3$, the matrix of which in the standard ordered basis is

$$\left[\begin{array}{ccc} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{array}\right].$$

Find a basis for the range of T and a basis for the null space of T.

Solution: The range is the column-space, which is the row-space of the following matrix (the transpose):

$$\left[\begin{array}{ccc} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 1 & 1 & 4 \end{array}\right]$$

which we can easily determine a basis of by putting it in row-reduced echelon form.

$$\left[\begin{array}{ccc} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 1 & 1 & 4 \end{array}\right] \rightarrow \left[\begin{array}{ccc} 1 & 0 & -1 \\ 0 & 1 & 5 \\ 0 & 1 & 5 \end{array}\right] \rightarrow \left[\begin{array}{ccc} 1 & 0 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{array}\right].$$

So a basis of the range is $\{(1,0,-1),(0,1,5)\}$.

The null space can be found by row-reducing the matrix

$$\left[\begin{array}{ccc} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{array}\right] \rightarrow \left[\begin{array}{ccc} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 5 & 5 \end{array}\right] \rightarrow \left[\begin{array}{ccc} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

So

$$\left\{\begin{array}{l} x - z = 0 \\ y + z = 0 \end{array}\right.$$

which implies

$$\left\{\begin{array}{l} x = z \\ y = -z \end{array}\right.$$

The solutions are parameterized by the one variable z, thus the null space has dimension equal to one. A basis is obtained by setting $z = 1$. Thus $\{(1,-1,1)\}$ is a basis for the null space.

Exercise 6: Let T be the linear operator on $\mathbb{R}^2$ defined by

$$T(x_1,x_2) = (-x_2,x_1).$$

(a) What is the matrix of T in the standard ordered basis for $\mathbb{R}^2$?

(b) What is the matrix of T in the ordered basis $\mathcal{B} = \{\alpha_1,\alpha_2\}$, where $\alpha_1 = (1,2)$ and $\alpha_2 = (1,-1)$?

(c) Prove that for every real number c the operator $(T - cI)$ is invertible.

(d) Prove that if $\mathcal{B}$ is any ordered basis for $\mathbb{R}^2$ and $[T]_{\mathcal{B}} = A$, then $A_{12}A_{21} \neq 0$.

Solution: (a) We must write $T\epsilon_1 = (0,1)$ and $T\epsilon_2 = (-1,0)$ in terms of ϵ_1 and ϵ_2. Clearly $T\epsilon_1 = \epsilon_2$ and $T\epsilon_2 = -\epsilon_1$. Thus the matrix is

$$\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right].$$

(b) We must write $T\alpha_1 = (-2, 1)$ and $T\alpha_2 = (1, 1)$ in terms of α_1, α_2. We can do this by row-reducing the augmented matrix

$$\left[\begin{array}{cc|cc} 1 & 1 & -2 & 1 \\ 2 & -1 & 1 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & 1 & -2 & 1 \\ 0 & -3 & 5 & -1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & 1 & -2 & 1 \\ 0 & 1 & -5/3 & 1/3 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -1/3 & 2/3 \\ 0 & 1 & -5/3 & 1/3 \end{array}\right]$$

Thus the matrix of T in the ordered basis $\mathcal{B}$ is

$$[T]_{\mathcal{B}} = \left[\begin{array}{cc} -1/3 & 2/3 \\ -5/3 & 1/3 \end{array}\right].$$

(c) The matrix of $T - cI$ with respect to the standard basis is

$$\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] - c\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] - \left[\begin{array}{cc} c & 0 \\ 0 & c \end{array}\right]$$

$$\left[\begin{array}{cc} -c & -1 \\ 1 & -c \end{array}\right].$$

Row-reducing the matrix

$$\left[\begin{array}{cc} -c & -1 \\ 1 & -c \end{array}\right] \rightarrow \left[\begin{array}{cc} 1 & -c \\ -c & -1 \end{array}\right] \rightarrow \left[\begin{array}{cc} 1 & -c \\ 0 & -1-c^2 \end{array}\right].$$

Now $-1 - c^2 \neq 0$ (since $c^2 \geq 0$). Thus we can continue row-reducing by dividing the second row by $-1 - c^2$ to get

$$\rightarrow \left[\begin{array}{cc} 1 & -c \\ 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right].$$

Thus the matrix has rank two, thus T is invertible.

(d) Let $\{\alpha_1, \alpha_2\}$ be any basis. Write $\alpha_1 = (a, b)$, $\alpha_2 = (c, d)$. Then $T\alpha_1 = (-b, a)$, $T\alpha_2 = (-d, c)$. We need to write $T\alpha_1$ and $T\alpha_2$ in terms of α_1 and α_2. We can do this by row reducing the augmented matrix

$$\left[\begin{array}{cc|cc} a & c & -b & -d \\ b & d & a & c \end{array}\right].$$

Since $\{\alpha_1, \alpha_2\}$ is a basis, the matrix $\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ is invertible. Thus (recalling Exercise 1.6.8, page 27), $ad - bc \neq 0$. Thus the matrix row-reduces to

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{ac+bd}{ad-bc} & \frac{c^2+d^2}{ad-bc} \\ 0 & 1 & \frac{a^2+b^2}{ad-bc} & \frac{ac+bd}{ad-bc} \end{array}\right].$$

Assuming $a \neq 0$ this can be shown as follows:

$$\rightarrow \left[\begin{array}{cc|cc} 1 & c/a & -b/a & -d/a \\ b & d & a & c \end{array}\right].$$

$$\to \left[\begin{array}{cc|cc} 1 & c/a & -b/a & -d/a \\ 0 & \frac{ad-bc}{a} & \frac{a^2+b^2}{a} & \frac{ac+bd}{a} \end{array}\right].$$

$$\to \left[\begin{array}{cc|cc} 1 & c/a & -b/a & -d/a \\ 0 & 1 & \frac{a^2+b^2}{ad-bc} & \frac{ac+bd}{ad-bc} \end{array}\right].$$

$$\to \left[\begin{array}{cc|cc} 1 & 0 & \frac{ac+bd}{ad-bc} & \frac{c^2+d^2}{ad-bc} \\ 0 & 1 & \frac{a^2+b^2}{ad-bc} & \frac{ac+bd}{ad-bc} \end{array}\right].$$

If $b \neq 0$ then a similar computation results in the same thing. Thus

$$[T]_{\mathcal{B}} = \left[\begin{array}{cc} \frac{ac+bd}{ad-bc} & \frac{c^2+d^2}{ad-bc} \\ \frac{a^2+b^2}{ad-bc} & \frac{ac+bd}{ad-bc} \end{array}\right].$$

Now $ad-bc \neq 0$ implies that at least one of a or b is non-zero and at least one of c or d is non-zero, it follows that $a^2+b^2>0$ and $c^2+d^2>0$. Thus $(a^2+b^2)(c^2+d^2) \neq 0$. Thus

$$\frac{a^2+b^2}{ad-bc} \cdot \frac{c^2+d^2}{ad-bc} \neq 0$$

Exercise 7: Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T(x_1,x_2,x_3) = (3x_1+x_3,\ \ -2x_1+x_2,\ \ -x_1+2x_2+4x_3).$$

(a) What is the matrix of T in the standard ordered basis for $\mathbb{R}^3$.

(b) What is the matrix of T in the ordered basis

$$(\alpha_1,\alpha_2,\alpha_3)$$

where $\alpha_1=(1,0,1)$, $\alpha_2=(-1,2,1)$, and $\alpha_3=(2,1,1)$?

(c) Prove that T is invertible and give a rule for T^{-1} like the one which defines T.

Solution: (a) As usual we can read the matrix in the standard basis right off the definition of T:

$$[T]_{\{\epsilon_1,\epsilon_2,\epsilon_3\}} = \left[\begin{array}{ccc} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{array}\right].$$

(b) $T\alpha_1 = (4,-2,3)$, $T\alpha_2 = (-2,4,9)$ and $T\alpha_3 = (7,-3,4)$. We must write these in terms of $\alpha_1,\alpha_2,\alpha_3$. We do this by row-reducing the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 4 & -2 & 7 \\ 0 & 2 & 1 & -2 & 4 & -3 \\ 1 & 1 & 1 & 3 & 9 & 4 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 4 & -2 & 7 \\ 0 & 2 & 1 & -2 & 4 & -3 \\ 0 & 2 & -1 & -1 & 11 & -3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 4 & -2 & 7 \\ 0 & 2 & 1 & -2 & 4 & -3 \\ 0 & 0 & -2 & 1 & 7 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 4 & -2 & 7 \\ 0 & 1 & 1/2 & -1 & 2 & -3/2 \\ 0 & 0 & 1 & -1/2 & -7/2 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 5/2 & 3 & 0 & 11/2 \\ 0 & 1 & 1/2 & -1 & 2 & -3/2 \\ 0 & 0 & 1 & -1/2 & -7/2 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 17/4 & 35/4 & 11/2 \\ 0 & 1 & 0 & -3/4 & 15/4 & -3/2 \\ 0 & 0 & 1 & -1/2 & -7/2 & 0 \end{array}\right]$$

Thus the matrix of T in the basis $\{\alpha_1, \alpha_2, \alpha_3\}$ is

$$[T]_{\{\alpha_1,\alpha_2,\alpha_3\}} = \left[\begin{array}{ccc} 17/4 & 35/4 & 11/2 \\ -3/4 & 15/4 & -3/2 \\ -1/2 & -7/2 & 0 \end{array}\right].$$

(c) We row reduce the augmented matrix (of T in the standard basis). If we achieve the identity matrix on the left of the dividing line then T is invertible and the matrix on the right will represent T^{-1} in the standard basis, from which we will be able read the rule for T^{-1} by inspection.

$$\left[\begin{array}{ccc|ccc} 3 & 0 & 1 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ -1 & 2 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} -1 & 2 & 4 & 0 & 0 & 1 \\ 3 & 0 & 1 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -2 & -4 & 0 & 0 & -1 \\ 3 & 0 & 1 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -2 & -4 & 0 & 0 & -1 \\ 0 & 6 & 13 & 1 & 0 & 3 \\ 0 & -3 & -8 & 0 & 1 & -2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -2 & -4 & 0 & 0 & -1 \\ 0 & 0 & -3 & 1 & 2 & -1 \\ 0 & -3 & -8 & 0 & 1 & -2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -2 & -4 & 0 & 0 & -1 \\ 0 & -3 & -8 & 0 & 1 & -2 \\ 0 & 0 & -3 & 1 & 2 & -1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -2 & -4 & 0 & 0 & -1 \\ 0 & 1 & 8/3 & 0 & -1/3 & 2/3 \\ 0 & 0 & -3 & 1 & 2 & -1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 4/3 & 0 & -2/3 & 1/3 \\ 0 & 1 & 8/3 & 0 & -1/3 & 2/3 \\ 0 & 0 & -3 & 1 & 2 & -1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 4/3 & 0 & -2/3 & 1/3 \\ 0 & 1 & 8/3 & 0 & -1/3 & 2/3 \\ 0 & 0 & 1 & -1/3 & -2/3 & 1/3 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4/9 & 2/9 & -1/9 \\ 0 & 1 & 0 & 8/9 & 13/9 & -2/9 \\ 0 & 0 & 1 & -1/3 & -2/3 & 1/3 \end{array}\right]$$

Thus T is invertible and the matrix for T^{-1} in the standard basis is

$$\left[\begin{array}{ccc} 4/9 & 2/9 & -1/9 \\ 8/9 & 13/9 & -2/9 \\ -1/3 & -2/3 & 1/3 \end{array}\right].$$

Thus $T^{-1}(x_1, x_2, x_3) = \left(\frac{4}{9}x_1 + \frac{2}{9}x_2 - \frac{1}{9}x_3,\ \ \frac{8}{9}x_1 + \frac{13}{9}x_2 - \frac{2}{9}x_3,\ \ -\frac{1}{3}x_1 - \frac{2}{3}x_2 + \frac{1}{3}x_3\right)$.

Exercise 8: Let θ be a real number. Prove that the following two matrices are similar over the field of complex numbers:

$$\left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right], \quad \left[\begin{array}{cc} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{array}\right]$$

(*Hint:* Let T be the linear operator on $\mathbb{C}^2$ which is represented by the first matrix in the standard ordered basis. Then find vectors α_1 and α_2 such that $T\alpha_1 = e^{i\theta}\alpha_1$, $T\alpha_2 = e^{-i\theta}\alpha_2$, and $\{\alpha_1, \alpha_2\}$ is a basis.)

Solution: Let $\mathcal{B}$ be the standard basis. Following the hint, let T be the linear operator on $\mathbb{C}^2$ which is represented by the first matrix in the standard ordered basis $\mathcal{B}$. Thus $[T]_{\mathcal{B}}$ is the first matrix above. Let $\alpha_1 = (i, 1)$, $\alpha_2 = (i, -1)$. Then α_1, α_2 are clealry linearly independent so $\mathcal{B}' = \{\alpha_1, \alpha_2\}$ is a basis for $\mathbb{C}^2$ (as a vector space over $\mathbb{C}$). Since $e^{i\theta} = \cos\theta + i\sin\theta$, it follows that $T\alpha_1 = (i\cos\theta - \sin\theta,\ \ i\sin\theta + \cos\theta) = (\cos\theta + i\sin\theta)(i, 1) = e^{i\theta}\alpha_1$ and similarly since and $e^{-i\theta} = \cos\theta - i\sin\theta$, it follows that $T\alpha_2 = e^{-i\theta}\alpha_2$. Thus the matrix of T with respect to $\mathcal{B}'$ is

$$[T]_{\mathcal{B}'} = \left[\begin{array}{cc} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{array}\right].$$

By Theorem 14, page 92, $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{B}'}$ are similar.

Exercise 9: Let V be a finite-dimensional vector space over the field F and let S and T be linear operators on V. We ask: When do there exist ordered bases $\mathcal{B}$ and $\mathcal{B}'$ for V such that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$? Prove that such bases exist if and only if there is an invertible linear operator U on V such that $T = USU^{-1}$. (*Outline of proof:* If $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$, let U be the operator which carries $\mathcal{B}$ onto $\mathcal{B}'$ and show that $S = UTU^{-1}$. Conversely, if $T = USU^{-1}$ for some invertible U, let $\mathcal{B}$ be any ordered basis for V and let $\mathcal{B}'$ be its image under U. Then show that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$.)

Solution: We follow the hint. Suppose there exist bases $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ and $\mathcal{B} = \{\beta_1, \dots, \beta_n\}$ such that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$. Let U be the operator which carries $\mathcal{B}$ onto $\mathcal{B}'$. Then by Theorem 14, page 92, $[USU^{-1}]_{\mathcal{B}'} = [U]_{\mathcal{B}}^{-1}[USU^{-1}]_{\mathcal{B}}[U]_{\mathcal{B}}$ and by the comments at the very bottom of page 90, this equals $[U]_{\mathcal{B}}^{-1}[U]_{\mathcal{B}}[S]_{\mathcal{B}}[U]_{\mathcal{B}}^{-1}[U]_{\mathcal{B}}$ which equals $[S]_{\mathcal{B}}$, which we've assumed equals $[T]_{\mathcal{B}'}$. Thus $[USU^{-1}]_{\mathcal{B}'} = [T]_{\mathcal{B}'}$. Thus $USU^{-1} = T$.

Conversely, assume $T = USU^{-1}$ for some invertible U. Let $\mathcal{B}$ be any ordered basis for V and let $\mathcal{B}'$ be its image under U. Then $[T]_{\mathcal{B}'} = [USU^{-1}]_{\mathcal{B}'} = [U]_{\mathcal{B}'}[S]_{\mathcal{B}'}[U]_{\mathcal{B}'}^{-1}$, which by Theorem 14, page 92, equals $[S]_{\mathcal{B}}$ (because U^{-1} carries $\mathcal{B}'$ into $\mathcal{B}$). Thus $[T]_{\mathcal{B}'} = [S]_{\mathcal{B}}$.

Exercise 10: We have seen that the linear operator T on $\mathbb{R}^2$ defined by $T(x_1, x_2) = (x_1, 0)$ is represented in the standard ordered basis by the matrix

$$A = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right].$$

This operator satisfies $T^2 = T$. Prove that if S is a linear operator on $\mathbb{R}^2$ such that $S^2 = S$, then $S = 0$, or $S = I$, or there is an ordered basis $\mathcal{B}$ for $\mathbb{R}^2$ such that $[S]_{\mathcal{B}} = A$ (above).

Solution: Suppose $S^2 = S$. Let ϵ_1, ϵ_2 be the standard basis vectors for $\mathbb{R}^2$. Consider $\{S\epsilon_1, S\epsilon_2\}$.

If both $S\epsilon_1 = S\epsilon_2 = 0$ then $S = 0$. Thus suppose WLOG that $S\epsilon_1 \neq 0$.

First note that if $x \in S(\mathbb{R}^2)$ then $x = S(y)$ for some $y \in \mathbb{R}^2$ and therefore $S(x) = S(S(y)) = S^2(y) = S(y) = x$. In other words $S(x) = x\ \forall\ x \in S(\mathbb{R}^2)$.

Case 1: Suppose $\exists\ c \in \mathbb{R}$ such that $S\epsilon_2 = cS\epsilon_1$. Then $S(\epsilon_2 - c\epsilon_1) = 0$. In this case S is singular because it maps a non-zero vector to zero. Thus since $S\epsilon_1 \neq 0$ we can conclude that $\dim(S(\mathbb{R}^2)) = 1$. Let α_1 be a basis for $S(\mathbb{R}^2)$. Let $\alpha_2 \in \mathbb{R}^2$ be such that $\{\alpha_1, \alpha_2\}$ is a basis for $\mathbb{R}^2$. Then $S\alpha_2 = k\alpha_1$ for some $k \in \mathbb{R}$. Let $\alpha_2' = \alpha_2 - k\alpha_1$. Then $\{\alpha_1, \alpha_2'\}$ span $\mathbb{R}^2$ because if $x = a\alpha_1 + b\alpha_2$ then $x = (a + bk)\alpha_1 + b\alpha_2'$. Thus $\{\alpha_1, \alpha_2'\}$ is a basis for $\mathbb{R}^2$. We now determine the matrix of S with respect to this basis. Since $\alpha_1 \in S(\mathbb{R}^2)$ and $S(x) = x\ \forall\ x \in S(\mathbb{R}^2)$, it follows that $S\alpha_1 = \alpha_1$. And consequently $S(\alpha_1) = 1 \cdot \alpha_1 + 0 \cdot \alpha_2'$. Thus the first column of the matrix of S with respect to α_1, α_2' is $[1, 0]^{\mathrm{T}}$. Also $S\alpha_2' = S(\alpha_2 - k\alpha_1) = S\alpha_2 - kS\alpha_1 = S\alpha_2 - k\alpha_1 = k\alpha_1 - k\alpha_1 = 0 = 0 \cdot \alpha_1 + 0 \cdot \alpha_2'$. So the second column of the matrix is $[0, 0]^{\mathrm{T}}$. Thus the matrix of S with respect to the basis $\{\alpha_1, \alpha_2'\}$ is exactly A.

Case 2: There does not exist $c \in \mathbb{R}$ such that $S\epsilon_2 = cS\epsilon_1$. In this case $S\epsilon_1$ and $S\epsilon_2$ are linearly independent from each other. Thus if we let $\alpha_i = S\epsilon_i$ then $\{\alpha_1, \alpha_2\}$ is a basis for $\mathbb{R}^2$. Now by assumption $S(x) = x\ \forall\ x \in S(\mathbb{R}^2)$, thus $S\alpha_1 = \alpha_1$ and $S\alpha_2 = \alpha_2$. Thus the matrix of S with respect to the basis $\{\alpha_1, \alpha_2\}$ is exactly the identity matrix I.

Exercise 11: Let W be the space of all $n \times 1$ column matrices over a field F. If A is an $n \times n$ matrix over F, then A defines a linear operator L_A on W through left multiplication: $L_A(X) = AX$. Prove that every linear operator on W is left multiplication by some $n \times n$ matrix, i.e., is L_A for some A.

Now suppose V is an n-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis for V. For each α in V, define $U\alpha = [\alpha]_{\mathcal{B}}$. Prove that U is an isomorphism of V onto W. If T is a linear operator on V, then UTU^{-1} is a linear operator on W. Accordingly, UTU^{-1} is left multiplication by some $n \times n$ matrix A. What is A?

Solution: *Part 1:* I'm confused by the first half of this question because isn't this exactly Theorem 11, page 87 in the special case $V = W$ where $\mathcal{B} = \mathcal{B}'$ is the standard basis of $F^{n\times 1}$. This special case is discussed on page 88 after Theorem 12, and in particular in Example 13. I don't know what we're supposed to add to that.

Part 2: Since $U(c\alpha_1 + \alpha_2) = [c\alpha_1 + \alpha_2]_{\mathcal{B}} = c[\alpha_1]_{\mathcal{B}} + [\alpha_2]_{\mathcal{B}} = cU(\alpha_1) + U(\alpha_2)$, U is linear, we just must show it is invertible. Suppose $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$. Let T be the function from W to V defined as follows:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \mapsto a_1\alpha_1 + \cdots a_n\alpha_n.$$

Then T is well defined and linear and it is also clear by inspection that TU is the identity transformation on V and UT is the identity transformation on W. Thus U is an isomorphism from V to W.

It remains to deterine the matrix of UTU^{-1}. Now $U\alpha_i$ is the standard $n \times 1$ matrix with all zeros except in the i-th place which equals one. Let $\mathcal{B}'$ be the standard basis for W. Then the matrix of U with respect to $\mathcal{B}$ and $\mathcal{B}'$ is the identity matrix. Likewise the matrix of U^{-1} with respect to $\mathcal{B}'$ and $\mathcal{B}$ is the identity matrix. Thus $[UTU^{-1}]_{\mathcal{B}'} = I[T]_{\mathcal{B}}I^{-1} = [T]_{\mathcal{B}}$. Therefore the matrix A is simply $[T]_{\mathcal{B}}$, the matrix of T with respect to $\mathcal{B}$.

Problem 12: Let V be an n-dimensional vector space over the field F, and let $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ be an ordered basis for V.

(a) According to Theorem 1, there is a unique linear operator T on V such that

$$T\alpha_j = \alpha_{j+1}, \quad j = 1, \dots, n-1, \quad T\alpha_n = 0.$$

What is the matrix A of T in the ordered basis $\mathcal{B}$.?

(b) Prove that $T^n = 0$ but $T^{n-1} \neq 0$.

(c) Let S be any linear operator on V such that $S^n = 0$ but $S^{n-1} \neq 0$. Prove that there is an ordered basis $\mathcal{B}'$ for V such that the matrix of S in the ordered basis $\mathcal{B}'$ is the matrix A of part (a).

(d) Prove that if M and N are $n\times n$ matrices over F such that $M^n = N^n = 0$ but $M^{n-1} \neq 0 \neq N^{n-1}$, then M and N are similar.

Solution: (a) The i-th column of A is given by the coefficients obtained by writing α_i in terms of $\{\alpha_1, \dots, \alpha_n\}$. Since $T\alpha_i = \alpha_{i+1}$, $i < n$ and $T\alpha_n = 0$, the matrix is therefore

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

(b) A has all zeros except 1's along the diagonal one below the main diagonal. Thus A^2 has all zeros except 1's along the diagonal that is two diagonals below the main diagonal, as follows:

$$A^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Similarly A^3 has all zeros except the diagonal three below the main diagonal. Continuing we see that A^{n-1} is the matrix that is all zeros except for the bottom left entry which is a 1:

$$A^{n-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Multiplying by A one more time then yields the zero matrix, $A^n = 0$. Since A represents T with respect to the basis $\mathcal{B}$, and A^i represents T^i, we see that $T^{n-1} \neq 0$ and $T^n = 0$.

(c) We will first show that $\dim(S^k(V)) = n - k$. Suppose $\dim(S(V)) = n$. Then $\dim(S^k(V)) = n \ \forall\ k = 1, 2, \dots,$ which contradicts the fact that $S^n = 0$. Thus it must be that $\dim(S(V)) \leq n - 1$. Now $\dim(S^2(V))$ cannot be greater than $\dim(S(V))$ because a linear transformation cannot map a space onto one with higher dimension. Thus $\dim(S^2(V)) \leq n - 1$. Suppose that $\dim(S^2(V)) = n - 1$. Thus $n - 1 = \dim(S^2(V)) \leq \dim(S(V)) \leq n - 1$. Thus it must be that $\dim(S(V)) = n - 1$. Thus S is an isomorphism on $S(V)$ because $S(V)$ and $S(S(V))$ have the same dimension. It follows that S^k is also an isomorphism on $S(V)\ \forall\ k \geq 2$. Thus it follows that $\dim(S^k(V)) = n - 1$ for all $k = 2, 3, 4, \dots,$ another contradiction. Thus $\dim(S^2(V)) \leq n - 2$. Suppose that $\dim(S^3(V)) = n - 2$, then it must be that $\dim(S^2(V)) = n - 2$ and therefore S is an isomorphism on $S^2(V)$, from which it follows that $\dim(S^k(V)) = n - 2$ for all $k = 3, 4, \dots,$ a contradiction. Thus $\dim(S^3(V)) \leq n - 3$. Continuing in this way we see that $\dim(S^k(V)) \leq n - k$. Thus $\dim(S^{n-1}(V)) \leq 1$. Since we are assuming $S^{n-1} \neq 0$ it follows that $\dim(S^{n-1}(V)) = 1$. We have seen that $\dim(S^k(V))$ cannot equal $\dim(S^{k+1}(V))$ for $k = 1, 2, \dots, n - 1$, thus it follows that the dimension must go down by one for each application of S. In other words $\dim(S^{n-2}(V))$ must equal 2, and then in turn $\dim(S^{n-3}(V))$ must equal 3, and generally $\dim(S^k(V)) = n - k$.

Now let α_1 be any basis vector for $S^{n-1}(V)$ which we have shown has dimension one. Now $S^{n-2}(V)$ has dimension two and S takes this space onto a space $S^{n-1}(V)$ of dimension one. Thus there must be $\alpha_2 \in S^{n-2}(V) \setminus S^{n-1}(V)$ such that $S(\alpha_2) = \alpha_1$. Since α_2 is not in the space generated by α_1 and $\{\alpha_1, \alpha_2\}$ are in the space $S^{n-2}(V)$ of dimension two, it follows that $\{\alpha_1, \alpha_2\}$ is a basis for $S^{n-2}(V)$. Now $S^{n-3}(V)$ has dimension three and S takes this space onto a space $S^{n-2}(V)$ of dimension two. Thus there must be $\alpha_3 \in S^{n-3}(V) \setminus S^{n-2}(V)$ such that $S(\alpha_3) = \alpha_2$. Since α_3 is not in the space generated by α_1 and α_2 and $\{\alpha_1, \alpha_2, \alpha_3\}$ are in the space $S^{n-3}(V)$ of dimension three, it follows that $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis for $S^{n-3}(V)$. Continuing in this way we produce a sequence of elements $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ that is a basis for $S^{n-k}(V)$ and such that $S(\alpha_i) = \alpha_{i-1}$ for all $i = 2, 3, \dots, k$. In particular we have a basis $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ for V and such that $S(\alpha_i) = \alpha_{i-1}$ for all $i = 2, 3, \dots, n$. Reverse the ordering of this bases to give $\mathcal{B} = \{\alpha_n, \alpha_{n-1}, \dots, \alpha_1\}$. Then $\mathcal{B}$ therefore is the required basis for which the matrix of S with respect to this basis will be the matrix given in part (a).

(d) Suppose S is the transformation of $F^{n\times 1}$ given by $v \mapsto Mv$ and similarly let T be the transformation $v \mapsto Nv$. Then $S^n = T^n = 0$ and $S^{n-1} \neq 0 \neq T^{n-1}$. Then we know from the previous parts of this problem that there is a basis $\mathcal{B}$ for which S is represented by the matrix from part (a). By Theorem 14, page 92, it follows that M is similar to the matrix in part (a). Likewise there's a basis $\mathcal{B}'$ for which T is represented by the matrix from part (a) and thus the matrix N is also similar to the matrix in part (a). Since similarity is an equivalence relation (see last paragraph page 94), it follows that since M and N are similar to the same matrix that they must be similar to each other.

Exercise 13: Let V and W be finite-dimensional vector spaces over the field F and let T be a linear transformation from V into W. If

$$\mathcal{B} = \{\alpha_1, \dots, \alpha_n\} \quad \text{and} \quad \mathcal{B}' = \{\beta_1, \dots, \beta_n\}$$

are ordered bases for V and W, respectively, define the linear transformations $E^{p,q}$ as in the proof of Theorem 5: $E^{p,q}(\alpha_i) = \delta_{i,q}\beta_p$. Then the $E^{p,q}$, $1 \leq p \leq m$, $1 \leq q \leq n$, form a basis for $L(V, W)$, and so

$$T = \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}E^{p,q}$$

for certain scalars A_{pq} (the coordinates of T in this basis for $L(V, W)$). Show that the matrix A with entries $A(p, q) = A_{pq}$ is precisely the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$.

Solution: Let $E^{p,q}_{\mathrm{M}}$ be the matrix of the linear transformation $E^{p,q}$ with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$. Then by the formula for a matrix associated to a linear transformation as given in the proof of Theorem 11, page 87, $E^{p,q}_{\mathrm{M}}$ is the matrix all of whose entries are zero except for the p, q-the entry which is one. Thus $A = \sum_{p,q} A_{p,q}E^{p,q}_{\mathrm{M}}$. Since the association between linear transformations and matrices is an isomorphism, $T \mapsto A$ implies $\sum_{p,q} A_{pq}E^{p,q} \mapsto \sum_{p,q} A_{pq}E^{p,q}_{\mathrm{M}}$. And thus A is exactly the matrix whose entries are the A_{pq}'s.

Section 3.5: Linear Functionals

Page 100: Typo line 5 from the top. It says $f(\alpha_i) = \alpha_i$, should be $f(\alpha_i) = a_i$.

Page 100: In Example 22, it says the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix}$$

is invertible "as a short computation shows." The way to see this is with what we know so far is to row reduce the matrix. As long as $t_1 \neq t_2$ we can get to

$$\begin{bmatrix} 1 & 0 & \frac{t_2-t_3}{t_2-t_1} \\ 0 & 1 & \frac{t_3-t_1}{t_2-t_1} \\ 0 & 0 & \frac{(t_3-t_1)(t_3-t_2)}{t_2^2-t_1^2} \end{bmatrix}.$$

Now we can continue and obtain
$$\left[\begin{array}{ccc} 1 & 0 & \frac{t_2-t_3}{t_2-t_1} \\ 0 & 1 & \frac{t_3-t_1}{t_2-t_1} \\ 0 & 0 & 1 \end{array}\right]$$
as long as $(t_3 - t_1)(t_3 - t_2) \neq 0$. From there we can finish row-reducing to obtain the identity. Thus we can row-reduce the matrix to the identity if and only if t_1, t_2, t_3 are distinct, that is no two of them are equal.

Exercise 1: In $\mathbb{R}^3$ let $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (0, 1, -2)$, $\alpha_3 = (-1, -1, 0)$.

(a) If f is a linear functional on $\mathbb{R}^3$ such that
$$f(\alpha_1) = 1, \quad f(\alpha_2) = -1, \quad f(\alpha_3) = 3,$$
and if $\alpha = (a, b, c)$, find $f(\alpha)$.

(b) Describe explicitly a linear functional f on $\mathbb{R}^3$ such that
$$f(\alpha_1) = f(\alpha_2) = 0 \quad \text{but} \quad f(\alpha_3) \neq 0.$$

(c) Let f be any linear functional such that
$$f(\alpha_1) = f(\alpha_2) = 0 \quad \text{and} \quad f(\alpha_3) \neq 0.$$
If $\alpha = (2, 3, -1)$, show that $f(\alpha) \neq 0$.

Solution: (a) We need to write (a, b, c) in terms of $\alpha_1, \alpha_2, \alpha_3$. We can do this by row reducing the following augmented matrix whose colums are the α_i's.
$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 0 & 1 & -1 & b \\ 1 & -2 & 0 & c \end{array}\right]$$
$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 0 & 1 & -1 & b \\ 0 & -2 & -1 & c-a \end{array}\right]$$
$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 0 & 1 & -1 & b \\ 0 & 0 & -1 & c-a+2b \end{array}\right]$$
$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 0 & 1 & -1 & b \\ 0 & 0 & 1 & a-2b-c \end{array}\right]$$
$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2a-2b-c \\ 0 & 1 & 0 & a-b-c \\ 0 & 0 & 1 & a-2b-c \end{array}\right]$$
Thus if $(a, b, c) = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$ then $x_1 = 2a - 2b - c$, $x_2 = a - b - c$ and $x_3 = a - 2b - c$. Now $f(a, b, c) = f(x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3) = x_1 f(\alpha_1) + x_2 f(\alpha_2) + x_3 f(\alpha_3) = (2a - 2b - c) \cdot 1 + (a - b - c) \cdot (-1) + (a - 2b - c) \cdot 3 = (2a - 2b - c) - (a - b - c) + (3a - 6b - 3c) = 4a - 7b - 3c$. In summary
$$f(\alpha) = 4a - 7b - 3c.$$

(b) Let $f(x, y, z) = x - 2y - z$. The $f(1, 0, 1) = 0$, $f(0, 1, -2) = 0$, and $f(-1, -1, 0) = 1$.

(c) Using part (a) we know that $\alpha = (2,3,-1) = -\alpha_1 - 3\alpha_3$ (plug in $a = 2$, $b = 3$, $c = -1$ for the formulas for x_1, x_2, x_3). Thus $f(\alpha) = -f(\alpha_1) - 3f(\alpha_3) = 0 - 3f(\alpha_3)$ and since $f(\alpha_3) \neq 0$, $-3f(\alpha_3) \neq 0$ and thus $f(\alpha) \neq 0$.

Exercise 2: Let $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ be the basis for $\mathbb{C}^3$ defined by

$$\alpha_1 = (1,0,-1), \quad \alpha_2 = (1,1,1), \quad \alpha_3 = (2,2,0).$$

Find the dual basis of $\mathcal{B}$.

Solution: The dual basis $\{f_1, f_2, f_3\}$ are given by $f_i(x_1,x_2,x_3) = \sum_{j=1}^{3} A_{ij}x_j$ where $(A_{1,1}, A_{1,2}, A_{1,3})$ is the solution to the system

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 0 \end{array}\right],$$

$(A_{2,1}, A_{2,2}, A_{2,3})$ is the solution to the system

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 0 \end{array}\right],$$

and $(A_{3,1}, A_{3,2}, A_{3,3})$ is the solution to the system

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{array}\right],$$

We row reduce the generic matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 1 & 1 & 1 & b \\ 2 & 2 & 0 & c \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & a+b-\frac{1}{2}c \\ 0 & 1 & 0 & c-b-a \\ 0 & 0 & 1 & b-\frac{1}{2}c \end{array}\right].$$

$a = 1, b = 0, c = 0 \Rightarrow f_1(x_1,x_2,x_3) = x_1 - x_2$
$a = 0, b = 1, c = 0 \Rightarrow f_2(x_1,x_2,x_3) = x_1 - x_2 + x_3$
$a = 0, b = 0, c = 1 \Rightarrow f_3(x_1,x_2,x_3) = -\frac{1}{2}x_1 + x_2 - \frac{1}{2}x_3.$

Then $\{f_1, f_2, f_3\}$ is the dual basis to $\{\alpha_1, \alpha_2, \alpha_3\}$.

Exercise 3: If A and B are $n \times n$ matrices over the field F, show that trace(AB) = trace(BA). Now show that similar matrices have the same trace.

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

$$\mathrm{trace}(AB) = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ik}B_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n} B_{ki}A_{ik} = \sum_{k=1}^{n}\sum_{i=1}^{n} B_{ki}A_{ik} = \sum_{k=1}^{n}(BA)_{kk} = \mathrm{trace}(BA).$$

Suppose A and B are similar. Then $\exists$ an invertible $n \times n$ matrix P such that $A = PBP^{-1}$. Thus trace(A) = trace(PBP^{-1}) = trace((P)(BP^{-1})) = trace((BP^{-1})(P)) = trace(B).

Exercise 4: Let V be the vector space of all polynomial functions p from R into R which have degree 2 or less:

$$p(x) = c_0 + c_1x + c_2x^2.$$

Define three linear functionals on V by

$$f_1(p) = \int_0^1 p(x)dx, \quad f_2(x) = \int_0^2 p(x)dx, \quad f_3(x) = \int_0^3 p(x)dx.$$

Show that $\{f_1, f_2, f_3\}$ is a basis for V^* by exhibiting the basis for V of which it is the dual.

Solution:

$$\int_0^a c_0 + c_1 x + c_2 x^2 dx$$
$$= c_0 x + \frac{1}{2} c_1 x^2 + \frac{1}{3} c_2 x^3 \,\Big|_0^a$$
$$= c_0 a + \frac{1}{2} c_1 a^2 + \frac{1}{3} c_2 a^3.$$

Thus

$$\int_0^1 p(x)dx = c_1 + \frac{1}{2} c_1 + \frac{1}{3} c_2$$
$$\int_0^2 p(x)dx = 2c_1 + 2c_1 + \frac{8}{3} c_2$$
$$\int_0^3 p(x)dx = 3c_1 + \frac{9}{2} c_1 + 9c_2$$

Thus we need to solve the following system three times

$$\begin{cases} c_1 + \frac{1}{2}c_1 + \frac{1}{3}c_2 = u \\ 2c_1 + 2c_1 + \frac{8}{3}c_2 = v \\ 3c_1 + \frac{9}{2}c_1 + 9c_2 = w \end{cases}$$

Once when $(u, v, w) = (1, 0, 0)$, once when $(u, v, w) = (0, 1, 0)$ and once when $(u, v, w) = (0, 0, 1)$.

We therefore row reduce the following matrix

$$\left[\begin{array}{ccc|ccc} 1 & 1/2 & 1/3 & 1 & 0 & 0 \\ 2 & 2 & 8/3 & 0 & 1 & 0 \\ 3 & 9/2 & 9 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 1/2 & 1/3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 3 & 8 & -3 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -2/3 & 2 & -1/2 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 2 & 3 & -3 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -2/3 & 2 & -1/2 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 3/2 & -3/2 & 1/2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -3/2 & 1/3 \\ 0 & 1 & 0 & -5 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -3/2 & 1/2 \end{array}\right].$$

Thus

$$\alpha_1 = 3 - 5x + \frac{3}{2}x^2$$
$$\alpha_2 = -\frac{3}{2} + 4x - \frac{3}{2}x^2$$
$$\alpha_3 = \frac{1}{3} - x + \frac{1}{2}x^2.$$

Exercise 5: If A and B are $n \times n$ complex matrices, show that $AB - BA = I$ is impossible.

Solution: Recall for $n \times n$ matrices M, $\text{trace}(M) = \sum_{i=1}^{n} M_{ii}$. The trace is clearly additive $\text{trace}(M_1 + M_2) = \text{trace}(M_1) + \text{trace}(M_2)$. We know from Exercise 3 that $\text{trace}(AB) = \text{trace}(BA)$. Thus $\text{trace}(AB - BA) = \text{trace}(AB) - \text{trace}(BA) = \text{trace}(AB) - \text{trace}(AB) = 0$. But $\text{trace}(I) = n$ and $n \neq 0$ in $\mathbb{C}$.

Exercise 6: Let m and n be positive integers and F a field. Let $f_1, \dots, f_m$ be linear functionals on F^n. For α in F^n define

$$T(\alpha) = (f_1(\alpha), \dots, f_m(\alpha)).$$

Show that T is a linear transformation from F^n into F^m. Then show that every linear transformation from F^n into F^m is of the above form, for some $f_1, \dots, f_m$.

Solution: Clearly T is a well defined function from F^n into F^m. We must just show it is linear. Let $\alpha, \beta \in F^n$, $c \in \mathbb{C}$. Then

$$T(c\alpha + \beta) = (f_1(c\alpha + \beta), \dots, f_m(c\alpha + \beta))$$

$$= (cf_1(\alpha) + f_1(\beta), \dots, cf_n(\alpha) + f_n(\beta))$$

$$= c(f_1(\alpha), \dots, f_n(\alpha)) + (f_1(\beta), \dots, f_n(\beta))$$

$$= cT(\alpha) + T(\beta).$$

Thus T is a linear transformation.

Let S be any linear transformation from F^n to F^m. Let M be the matrix of S with respect to the standard bases of F^n and F^m. Then M is an $m \times n$ matrix and S is given by $X \mapsto MX$ where we identify F^n as $F^{n\times 1}$ and F^m with $F^{m\times 1}$. Now for each $i = 1, \dots, m$ let $f_i(x_1, \dots, x_n) = \sum_{j=1}^{n} M_{ij}x_j$. Then $X \mapsto MX$ is the same as $X \mapsto (f_1(X), \dots, f_m(x))$ (keeping in mind our identification of F^m with $F^{m\times 1}$). Thus S has been written in the desired form.

Exercise 7: Let $\alpha_1 = (1, 0, -1, 2)$ and $\alpha_2 = (2, 3, 1, 1)$, and let W be the subspace of $\mathbb{R}^4$ spanned by α_1 and α_2. Which linear functionals f:

$$f(x_1, x_2, x_3, x_4) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$

are in the annihilator of W?

Solution: The two vectors α_1 and α_2 are linearly independent since neither is a multiple of the other. Thus W has dimension 2 and $\{\alpha_1, \alpha_2\}$ is a basis for W. Therefore a functional f is in the annihilator of W if and only if $f(\alpha_1) = f(\alpha_2) = 0$. We find such f by solving the system

$$\begin{cases} f(\alpha_1) = 0 \\ f(\alpha_2) = 0 \end{cases}$$

or equivalently

$$\begin{cases} c_1 - c_3 + 2c_4 = 0 \\ 2c_1 + 3c_2 + c_3 + c_4 = 0 \end{cases}$$

We do this by row reducing the matrix

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

$$\to \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

Therefore

$$c_1 = c_3 - 2c_4$$

$$c_2 = -c_3 + c_4.$$

The general element of W^0 is therefore

$$f(x_1, x_2, x_3, x_4) = (c_3 - 2c_4)x_1 + (c_3 + c_4)x_2 + c_3x_3 + c_4x_4,$$

for arbitrary elements c_3 and c_4. Thus W^0 has dimension 2 as expected.

Exercise 8: Let W be the subspace of $\mathbb{R}^5$ which is spanned by the vectors

$$\alpha_1 = \epsilon_1 + 2\epsilon_2 + \epsilon_3, \quad \alpha_2 = \epsilon_2 + 3\epsilon_3 + 3\epsilon_4 + \epsilon_5$$

$$\alpha_3 = \epsilon_1 + 4\epsilon_2 + 6\epsilon_3 + 4\epsilon_4 + \epsilon_5.$$

Find a basis for W^0.

Solution: The vectors $\alpha_1, \alpha_2, \alpha_3$ are linearly independent as can be seen by row reducing the matrix

$$\begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 3 & 1 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 3 & 1 \\ 0 & 2 & 5 & 4 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & -5 & -6 & -2 \\ 0 & 1 & 3 & 3 & 1 \\ 0 & 0 & -1 & -2 & -1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & -5 & -6 & -2 \\ 0 & 1 & 3 & 3 & 1 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & 4 & 3 \\ 0 & 1 & 0 & -3 & -2 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix}.$$

Thus W has dimension 3 and $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis for W. We know every functional is given by $f(x_1, x_2, x_3, x_4, x_5) = c_1x_2 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5$ for some $c_1, \ldots, c_5$. From the row reduced matrix we see that the general solution for an element of W^0 is

$$f(x_1, x_2, x_3, x_4, x_5) = (-4c_4 - 3c_5)x_1 + (3c_4 + 2c_5)x_2 - (2c_4 + c_5)x_3 + c_4x_4 + c_5x_5.$$

Exercise 9: Let V be the vector space of all 2×2 matrices over the field of real numbers, and let

$$B = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}.$$

Let W be the subspace of V consisting of all A such that $AB = 0$. Let f be a linear functional on V which is in the annihilator of W. Suppose that $f(I) = 0$ and $f(C) = 3$, where I is the 2×2 identity matrix and

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Find $f(B)$.

Solution: The general linear functional on V is of the form $f(A) = aA_{11} + bA_{12} + cA_{21} + dA_{22}$ for some $a, b, c, d \in \mathbb{R}$. If $A \in W$ then

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

implies $y = 2x$ and $w = 2y$. So W consists of all matrices of the form

$$\begin{bmatrix} x & 2x \\ y & 2y \end{bmatrix}$$

Now $f \in W^0 \Rightarrow f\left(\begin{bmatrix} x & 2x \\ y & 2y \end{bmatrix}\right) = 0 \ \forall \ x, y \in \mathbb{R} \Rightarrow ax + 2bx + cy + 2dy = 0 \ \forall \ x, y \in \mathbb{R} \Rightarrow (a + 2b)x + (c + 2d)y = 0 \ \forall \ x, y \in \mathbb{R}$ $\Rightarrow b = -\frac{1}{2}a$ and $d = -\frac{1}{2}c$. So the general $f \in W^0$ is of the form

$$f(A) = aA_{11} - \frac{1}{2}aA_{12} + cA_{21} - \frac{1}{2}cA_{22}.$$

Now $f(C) = 3 \Rightarrow d = 3 \Rightarrow -\frac{1}{2}c = 3 \Rightarrow c = -6$. And $f(I) = 0 \Rightarrow a - \frac{1}{2}c = 0 \Rightarrow c = 2a \Rightarrow a = -3$. Thus

$$f(A) = -3A_{11} + \frac{3}{2}A_{12} - 6A_{21} + 3A_{22}.$$

Thus

$$f(B) = -3 \cdot 2 + \frac{3}{2} \cdot (-2) - 6 \cdot (-1) + 3 \cdot 1 = 0.$$

Exercise 10: Let F be a subfield of the complex numbers. We define n linear functionals on F^n ($n \geq 2$) by

$$f_k(x_1, \ldots, x_n) = \sum_{j=1}^{n} (k - j)x_j, \quad 1 \leq k \leq n.$$

What is the dimension of the subspace annihilated by $f_1, \ldots, f_n$?

Solution: N_{f_k} is the subspace annihilated by f_k. By the comments on page 101, N_{f_k} has dimension $n - 1$. Now the standard basis vector ϵ_2 is in N_{f_2} but is not in N_{f_1}. Thus N_{f_1} and N_{f_2} are distinct hyperspaces. Thus their intersection has dimension $n - 2$. Now ϵ_3 is in N_{f_3} but is not in $N_{f_1} \cup N_{f_2}$. Thus $N_{f_1} \cap N_{f_2} \cap N_{f_3}$ is the intersection of three distinct hyperspaces and so has dimension $n - 3$. Continuing in this way, $\epsilon_i \notin \cup_{j=1}^{i-1} N_{f_i}$. Thus $\cup_{j=1}^{i} N_{f_i}$ is the intersection of i distinct hyperspaces and so has dimension $n - i$. Thus when $i = n$ we have $\cup_{j=1}^{n} N_{f_i}$ has dimension 0.

Exercise 11: Let W_1 and W_2 be subspace of a finite-dimensional vector space V.

(a) Prove that $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$.

(b) Prove that $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

Solution: (a) $f \in (W_1 + W_2)^0 \Rightarrow f(v) = 0 \ \forall \ v \in W_1 + W_2 \Rightarrow f(w_1 + w_2) = 0 \ \forall \ w_1 \in W_1, w_2 \in W_2 \Rightarrow f(w_1) = 0 \ \forall \ w_1 \in W_1$ (take $w_2 = 0$) and $f(w_2) = 0 \ \forall \ w_2 \in W_2$ (take $w_1 = 0$). Thus $f \in W_1^0$ and $f \in W_2^0$. Thus $f \in W_1^0 \cap W_2^0$. Thus $(W_1 + W_2)^0 \subseteq W_1^0 \cap W_2^0$.

Conversely, let $f \in W_1^0 \cap W_2^0$. Let $v \in W_1 + W_2$. Then $v = w_1 + w_2$ where $w_i \in W_i$. Thus $f(v) = f(w_1 + w_2) = f(w_1) + f(w_2) = 0 + 0$ (since $f \in W_1^0$ and $f \in W_2^0$). Thus $f(v) = 0 \ \forall \ v \in W_1 + W_2$. Thus $f \in (W_1 + W_2)^0$. Thus $W_1^0 \cap W_2^0 \subseteq (W_1 + W_2)^0$.

Since $(W_1 + W_2)^0 \subseteq W_1^0 \cap W_2^0$ and $W_1^0 \cap W_2^0 \subseteq (W_1 + W_2)^0$ it follows that $W_1^0 \cap W_2^0 = (W_1 + W_2)^0$

(b) $f \in W_1^0 + W_2^0 \Rightarrow f = f_1 + f_2$, for some $f_i \in W_i^0$. Now let $v \in W_1 \cap W_2$. Then $f(v) = (f_1 + f_2)(v) = f_1(v) + f_2(v) = 0 + 0$. Thus $f \in (W_1 \cap W_2)^0$. Thus $W_1^0 + W_2^0 \subseteq (W_1 \cap W_2)^0$.

Now let $f \in (W_1 \cap W_2)^0$. In the proof of Theorem 6 on page 46 it was shown that we can choose a basis for $W_1 + W_2$

$$\{\alpha_1, \dots, \alpha_k, \quad \beta_1, \dots, \beta_m, \quad \gamma_1, \dots, \gamma_n\}$$

where $\{\alpha_1, \dots, \alpha_k\}$ is a basis for $W_1 \cap W_2$, $\{\alpha_1, \dots, \alpha_k, \quad \beta_1, \dots, \beta_m\}$ is a basis for W_1 and $\{\alpha_1, \dots, \alpha_k, \quad \gamma_1, \dots, \gamma_n\}$ is a basis for W_2. We expand this to a basis for all of V

$$\{\alpha_1, \dots, \alpha_k, \quad \beta_1, \dots, \beta_m, \quad \gamma_1, \dots, \gamma_n, \quad \lambda_1, \dots, \lambda_\ell\}.$$

Now the general element $v \in V$ can be written as

$$v = \sum_{i=1}^{k} x_i\alpha_i + \sum_{i=1}^{m} y_i\beta_i + \sum_{i=1}^{n} z_i\gamma_i + \sum_{i=1}^{\ell} w_i\lambda_i \tag{20}$$

and f is given by

$$f(v) = \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{m} b_i y_i + \sum_{i=1}^{n} c_i z_i + \sum_{i=1}^{\ell} d_i w_i$$

for some constants a_i, b_i, c_i, d_i. Since $f(v) = 0$ for all $v \in W_1 \cap W_2$, it follows that $a_1 = \cdots = a_k = 0$. So

$$f(v) = \sum b_i y_i + \sum c_i z_i + \sum d_i w_i.$$

Define

$$f_1(v) = \sum c_i z_i + \sum d_i w_i$$

and

$$f_2(v) = \sum b_i y_i.$$

Then $f = f_1 + f_2$. Now if $v \in W_1$ then

$$v = \sum_{i=1}^{k} x_i\alpha_i + \sum_{i=1}^{m} y_i\beta_i$$

so that the coefficients z_i and w_i in (20) are all zero. Thus $f_1(v) = 0$. Thus $f_1 \in W_1^0$. Similarly if $v \in W_2$ then the coefficients y_i and w_i in (20) are all zero and thus $f_2(v) = 0$. So $f_2 \in W_2$. Thus $f = f_1 + f_2$ where $f_1 \in W_1^0$ and $f_2 \in W_2^0$. Thus $f \in W_1^0 + W_2^0$. Thus $(W_1 \cap W_2)^0 \subseteq W_1^0 + W_2^0$.

Thus $(W_1 \cap W_2)^0 \subseteq W_1^0 + W_2^0$.

Exercise 12: Let V be a finite-dimensional vector space over the field F and let W be a subspace of V. If f is a linear functional on W, prove that there is a linear functional g on V suvch that $g(\alpha) = f(\alpha)$ for each α in the subspace W.

Solution: Let $\mathcal{B}$ be a basis for W and let $\mathcal{B}'$ be a basis for V such that $\mathcal{B} \subseteq \mathcal{B}'$. A linear function on a vector space is uniquely determined by its values on a basis, and conversely any function on the basis can be extended to a linear function on the space. Thus we define g on $\mathcal{B}$ by $g(\beta) = f(\beta)\ \forall\, \beta \in \mathcal{B}$. Then define $g(\beta) = 0$ for all $\beta \in \mathcal{B}' \setminus \mathcal{B}$. Since we have defined g on $\mathcal{B}'$ it defines a linear functional on V and since it agrees with f on a basis for W it agrees with f on all of W.

Exercise 13: Let F be a subfield of the field of complex numbers and let V be any vector space over F. Suppose that f and g are linear functionals on V such that the function h defined by $h(\alpha) = f(\alpha)g(\alpha)$ is also a linear functional on V. Prove that either $f = 0$ or $g = 0$.

Solution: Suppose neither f nor g is the zero function. We will derive a contradiction. Let $v \in V$. Then $h(2v) = f(2v)g(2v) = 4f(v)g(v)$. But also $h(2v) = 2h(v) = 2f(v)g(v)$. Therefore $f(v)g(v) = 2f(v)g(v)\ \forall\ v \in V$. Thus $f(v)g(v) = 0\ \forall\ v \in V$. Let $\mathcal{B}$ be a basis for V. Let $\mathcal{B}_1 = \{\beta \in \mathcal{B} \mid f(\beta) = 0\}$ and $\mathcal{B}_2 = \{\beta \in \mathcal{B} \mid g(\beta) = 0\}$. Since $f(\beta)g(\beta) = 0\ \forall\ \beta \in \mathcal{B}$, we have $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$. Suppose $\mathcal{B}_1 \subseteq \mathcal{B}_2$. Then $\mathcal{B}_2 = \mathcal{B}$ and consequently g is the zero function. Thus $\mathcal{B}_1 \not\subseteq \mathcal{B}_2$. And similarly $\mathcal{B}_2 \not\subseteq \mathcal{B}_1$. Thus we can choose $\beta_1 \in \mathcal{B}_1 \setminus \mathcal{B}_2$ and $\beta_2 \in \mathcal{B}_2 \setminus \mathcal{B}_1$. So we have $f(\beta_2) \neq 0$ and $g(\beta_1) \neq 0$. Then $f(\beta_1+\beta_2)g(\beta_1+\beta_2) = f(\beta_1)g(\beta_1) + f(\beta_2)g(\beta_1) + f(\beta_1)g(\beta_2) + f(\beta_2)g(\beta_2)$. Since $f(\beta_1) = g(\beta_2) = 0$, this equals $f(\beta_2)g(\beta_1)$ which is non-zero since each term is non-zero. And this contradicts the fact that $f(v)g(v) = 0\ \forall\ v \in V$.

Exercise 14: Let F be a field of characteristic zero and let V be a finite-dimensional vector space over F. If $\alpha_1, \ldots, \alpha_m$ are finitely many vectors in V, each different from the zero vector, prove that there is a linear functional f on V such that

$$f(\alpha_i) \neq 0, \quad i = 1, \ldots, m.$$

Solution: Re-index if necessary so that $\{\alpha_1, \ldots, \alpha_k\}$ is a basis for the subspace generated by $\{\alpha_1, \ldots, \alpha_m\}$. So each $\alpha_{k+1}, \ldots, \alpha_m$ can be written in terms of $\alpha_1, \ldots, \alpha_k$. Extend $\{\alpha_1, \ldots, \alpha_k\}$ to a basis for V

$$\{\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_n\}.$$

For each $i = k+1, \ldots, m$ write $\alpha_i = \sum_{j=1}^{k} A_{ij}\alpha_j$. Since $\alpha_{k+1}, \ldots, \alpha_m$ are all non-zero, for each $i = k+1, \ldots, m\ \exists\ j_i \leq k$ such that $A_{ij_i} \neq 0$. Now define f by mapping $\alpha_1, \ldots, \alpha_k$ to k arbitrary non-zero values and map β_i to zero $\forall\ i$. Then $f(\alpha_{k+1}) = \sum_{j=1}^{k} A_{k+1,j} f(\alpha_j)$. If $f(\alpha_{k+1}) = 0$ then leaving $f(\alpha_i)$ fixed for all $i \leq k$ and adjusting $f(\alpha_{j_{k+1}})$, it equals zero for exactly one possible value of $f(\alpha_{j_{k+1}})$ (since $A_{k+1,j_{k+1}} \neq 0$). Thus we can redefine $f(\alpha_{j_{k+1}})$ so that $f(\alpha_{k+1}) \neq 0$ while maintaining $f(\alpha_{j_{k+1}}) \neq 0$.

Now if $f(\alpha_{k+2}) = 0$, then leaving $f(\alpha_i)$ fixed for $i \neq j_{k+2}$, it equals zero for exactly one possible value of $f(\alpha_{j_{k+2}})$ (since $A_{k+2,j_{k+2}} \neq 0$) So we can adjust $f(\alpha_{j_{k+2}})$ so that $f(\alpha_{k+2}) \neq 0$ and $f(\alpha_{k+1}) \neq 0$ and $f(\alpha_{k+2}) \neq 0$ simultaneously.

Continuing in this way we can adjust $f(\alpha_{j_{k+3}}), \ldots, f(\alpha_{j_m})$ as necessary until all $f(\alpha_{k+1}), \ldots, f(\alpha_m)$ are non-zero and also all of $f(\alpha_1), \ldots, f(\alpha_k)$ are non-zero.

Exercise 15: According to Exercise 3, similar matrices have the same trace. Thus we can define the trace of a linear operator on a finite-dimensional space to be the trace of any matrix which represents the operator in an ordered basis. This is well-defined since all such representing matrices for one operator are similar.

Now let V be the space of all 2×2 matrices over the field F and let P be a fixed 2×2 matrix. Let T be the linear operator on V defined by $T(A) = PA$. Prove that trace$(T) = 2$trace(P).

Solution: Write

$$P = \left[\begin{array}{cc} P_{11} & P_{12} \\ P_{21} & P_{22} \end{array}\right].$$

Let

$$e_{11} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right], \quad e_{12} = \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]$$

$$e_{21} = \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right], \quad e_{22} = \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$$

Then $\mathcal{B} = \{e_{11}, e_{12}, e_{21}, e_{22}\}$ is an ordered basis for V. We find the matrix of the linear transformation with respect to this basis.

$$T(e_{11}) = \left[\begin{array}{cc} P_{11} & 0 \\ P_{21} & 0 \end{array}\right] = P_{11}e_{11} + P_{21}e_{21}$$

$$T(e_{12}) = \left[\begin{array}{cc} 0 & P_{11} \\ 0 & P_{21} \end{array}\right] = P_{11}e_{12} + P_{21}e_{22}$$

$$T(e_{21}) = \left[\begin{array}{cc} P_{21} & 0 \\ P_{22} & 0 \end{array}\right] = P_{12}e_{11} + P_{22}e_{21}$$

$$T(e_{22}) = \left[\begin{array}{cc} 0 & P_{12} \\ 0 & P_{22} \end{array}\right] = P_{12}e_{12} + P_{22}e_{22}.$$

Thus the matrix of T with respect to $\mathcal{B}$ is

$$\left[\begin{array}{cccc} P_{11} & 0 & P_{12} & 0 \\ 0 & P_{11} & 0 & P_{12} \\ P_{21} & 0 & P_{22} & 0 \\ 0 & P_{21} & 0 & P_{22} \end{array}\right].$$

The trace of this matrix is $2P_{11} + 2P_{22} = 2\text{trace}(P)$.

Exercise 16: Show that the trace functional on $n \times n$ matrices is unique in the following sense. If W is the space of $n \times n$ matrices over the field F and if f is a linear functional on W such that $f(AB) = f(BA)$ for each A and B in W, then f is a scalar multiple of the trace function. If, in addition, $f(I) = n$ then f is the trace function.

Solution: Let A and B be $n \times n$ matrices. The ℓ, m entry in AB is

$$(AB)_{\ell m} = \sum_{k=1}^{n} A_{\ell k} B_{km} \tag{21}$$

and the ℓ, m entry in BA is

$$(BA)_{\ell m} = \sum_{k=1}^{n} B_{\ell k} A_{km}. \tag{22}$$

Fix $i, j \in \{1, \ldots, n\}$ such that $i > j$. Let A be the matrix where $A_{ij} = 1$ and all other entries are zero. Let B be the matrix where $B_{ii} = 1$ and all other entries are zero. Consider the general element of AB

$$(AB)_{\ell m} = \sum_{k=1}^{n} A_{\ell k} B_{km}.$$

The only non-zero A in the sum on the right is A_{ij}. But $B_{jm} = 0$ since $j > i$ and only $B_{ii} \neq 0$. Thus AB is the zero matrix.

Now we compute BA. From (22) the only non-zero term is when $\ell = i$, $m = j$ and $k = i$.

Thus the matrix AB has zeros in every position except for the i, j position where it equals one.

Now the general functional on $n \times n$ matrices is of the form

$$f(M) = \sum_{\ell=1}^{n} \sum_{m=1}^{n} c_{\ell m} M_{\ell m}$$

for some constants $c_{\ell m}$. Now $f(AB) = f(0) = 0$ and $f(BA) = c_{ij}$. So if $f(AB) = f(BA)$ then it follows that $c_{ij} = 0$.

Thus we have shown that $c_{ij} = 0$ for all $i > j$. Similarly $c_{ij} = 0$ for all $i < j$. Thus the only possible non-zero coefficients are $c_{11}, \ldots, c_{nn}$.

$$f(M) = \sum_{i=1}^{n} c_{ii} M_{ii}.$$

We will be done if we show $c_{11} = c_{mm}$ for all $m = 2, \ldots, n$. Fix $2 \leq i \leq n$. Let A be the matrix such that $A_{11} = A_{i1} = 1$ and $A_{\ell m} = 0$ in all other positions. Let $B = A^{\mathrm{T}}$. Then AB is zero in every position except $A_{11} = A_{1i} = A_{i1} = A_{ii} = 1$. And BA is zero in every position except $(BA)_{11} = 2$. Thus $f(AB) = c_{11} + c_{ii}$ and $f(BA) = 2c_{11}$. Thus if $f(AB) = f(BA)$ then $c_{11} + c_{ii} = 2c_{11}$ which implies $c_{11} = c_{ii}$. Thus there's a constant c such that $c_{ii} = c$ for all i.

Thus f is given by

$$f(M) = \sum_{k=1}^{n} c M_{ii}.$$

If $f(I) = n$ then $c = 1$ and we have the trace function.

Exercise 17: Let W be the space of $n \times n$ matrices over the field F, and let W_0 be the subspace spanned by the matrices C of the form $C = AB - BA$. Prove that W_0 is exactly the subspace of matrices which have trace zero. (*Hint:* What is the dimension of the space of matrices of trace zero? Use the matrix 'units,' i.e., matrices with exactly one non-zero entry, to construct enough linearly independent matrices of the form $AB - BA$.)

Solution: Let $W' = \{w \in W \mid \text{trace}(w) = 0\}$. We want to show $W' = W_0$. We know from Exercise 3 that $\text{trace}(AB - BA) = 0$ for all matrices A, B. Since matrices of the form $AB - BA$ span W_0, it follows that $\text{trace}(M) = 0$ for all $M \in W_0$. Thus $W_0 \subseteq W'$.

Since the trace function is a linear functional, the dimension of W' is $\dim(W) - 1 = n^2 - 1$. Thus if we show the dimension of W_0 is also $n^2 - 1$ then we will be done. We do this by exhibiting $n^2 - 1$ linearly independent elements of W_0. Denote by E_{ij} the matrix with a one in the i, j position and zeros in all other positions. Let $H_{ij} = E_{ii} - E_{jj}$. Let $\mathcal{B} = \{E_{ij} \mid i \neq j\} \cup \{H_{1,i} \mid 2 \leq i \leq n\}$. We will show that $\mathcal{B} \subseteq W_0$ and that $\mathcal{B}$ is a linearly independent set. First, it clear that they are linearly independent because E_{ij} is the only vector in $\mathcal{B}$ with a non-zero value in the i, j position and $H_{1,i}$ is the only vector in $\mathcal{B}$ with a non-zero value in the i, i position. Now $2E_{ij} = H_{ij}E_{ij} - E_{ij}H_{ij}$ and $H_{ij} = E_{ij}E_{ji} - E_{ji}E_{ij}$. Thus $E_{ij} \in W_0$ and $H_{ij} \in W_0$. Now $|\mathcal{B}| = |\{E_{ij} \mid i \neq j\}| + |\{H_{1,i} \mid 2 \leq i \leq n\}| = (n^2 - n) + (n - 1) = n^2 - 1$ Thus we are done.

Section 3.6: The Double Dual

Exercise 1: Let n be a positive integer and F a field. Let W be the set of all vectors $(x_1, \ldots, x_n)$ in F^n such that $x_1 + \cdots + x_n = 0$.

(a) Prove that W^0 consists of all linear functionals f of the form

$$f(x_1, \ldots, x_n) = c \sum_{j=1}^{n} x_j.$$

(b) Show that the dual space W^* of W can be 'naturally' identified with the linear functionals

$$f(x_1, \ldots, x_n) = c_1 x_1 + \cdots + c_n x_n$$

on F^n which satisfy $c_1 + \cdots + c_n = 0$.

Solution: (a) Let g be the functional $g(x_1, \ldots, x_n) = x_1 + \cdots + x_n$. Then W is exactly the kernel of g. Thus $\dim(W) = n - 1$. Let $\alpha_i = \epsilon_1 - \epsilon_{i+1}$ for $i = 1, \ldots, n - 1$. Then $\{\alpha_1, \ldots, \alpha_{n-1}\}$ are linearly independent and are all in W so they must be a basis for W. Let $f(x_1, \ldots, x_n) = c_1 x_1 + \cdots + c_n x_n$ be a linear functional. Then $f \in W^0 \Rightarrow f(\alpha_1) = \cdots = f(\alpha_n) = 0 \Rightarrow c_1 - c_i = 0\ \forall$ $i = 2, \ldots, n \Rightarrow \exists\, c$ such that $c_i = c\ \forall\ i$. Thus $f(x_1, \ldots, x_n) = c(x_1 + \cdots + x_n)$.

(b) Consider the sequence of functions

$$W \to (F^n)^* \to W^*$$

where the first function is

$$(c_1, \ldots, c_n) \mapsto f_{c_1,\ldots,c_n}$$

where $f_{c_1,\ldots,c_n}(x_1, \ldots, x_n) = c_1 x_1 + \cdots c_n x_n$ and the second function is restriction from F^n to W. We know both W and W^* have the same dimension. Thus if we show the composition of these two functions is one-to-one then it must be an isomorphism. Suppose $(c_1, \ldots, c_n) \in W \mapsto f_{c_1,\ldots,c_n} = 0 \in W^*$.

Then $\sum c_i = 0$ and $\sum c_i x_i = 0$ for all $(x_1, \ldots, x_n) \in W$.

In other words $\sum c_i = 0$ and $\sum c_i x_i = 0$ for all $(x_1, \ldots, x_n)$ such that $\sum x_i = 0$.

Let $\{\alpha_1, \dots, \alpha_{n-1}\}$ be the basis for W from part (a). Then $f_{c_1,\dots,c_n}(\alpha_i) = 0\ \forall\ i = 1, \dots, n-1$; which implies $c_1 = c_i\ \forall\ i = 2, \dots, n$. Thus $\sum c_i = (n-1)c_1$. But $\sum c_i = 0$, thus $c_1 = 0$. Thus $f_{c_1,\dots,c_n}$ is the zero function.

Thus the mapping $W \to W^*$ is a natural isomorphism. We therefore naturally identify each element in W^* with a linear functional $f(x_1, \dots, x_n) = c_1x_1 + \cdots c_nx_n$ where $\sum c_i = 0$.

Exercise 2: Use Theorem 20 to prove the following. If W is a subspace of a finite-dimensional vector space V and if $\{g_1, \dots, g_r\}$ is any basis for W^0, then

$$W = \bigcap_{i=1}^{r} N_{g_i}.$$

Solution: $\{g_1, \dots, g_r\}$ a basis for $W^0 \Rightarrow g_i \in W^0\ \forall\ i \Rightarrow g_i(W) = 0\ \forall\ i \Rightarrow W \subseteq N_{g_i}\ \forall\ i \Rightarrow W \subseteq \bigcap_{i=1}^r N_{g_i}$. Let $n = \dim(V)$. By Theorem 2, page 71, we know $\dim(N_{g_1}) = n-1$. Since the g_i's are linearly independent, g_2 is not a multiple of g_1, thus by Theorem 20, $N_{g_1} \not\subseteq N_{g_2}$. Thus $\dim(N_{g_1} \cap N_{g_2}) \leq n-2$. By Theorem 20 again, $N_{g_3} \not\subseteq N_{g_1} \cap N_{g_2}$ since g_3 is not a linear combination of g_1 and g_2. Thus $\dim(N_{g_1} \cap N_{g_2} \cap N_{g_3}) \leq n-3$. By induction $\dim(\bigcap_{i=1}^r N_{g_i}) \leq n-r$. Now by Theorem 16, $\dim(W) = n-r$. Thus since $W \subseteq \bigcap_{i=1}^r N_{g_i}$, it follows that $\dim(\bigcap_{i=1}^r N_{g_i}) \geq n-r$. Thus it must be that $\dim(\bigcap_{i=1}^r N_{g_i}) = n-r$ and it must be that $W = \bigcap_{i=1}^r N_{g_i}$ since we have shown the left hand side is contained in the right hand side and both sides have the same dimension.

Exercise 3: Let S be a set, F a field, and $V(S; F)$ the space of all functions from S into F:

$$(f+g)(x) = f(x) + g(x)$$

$$(cf)(x) = cf(x).$$

Let W be any n-dimensional subspace of $V(S; F)$. Show that there exist points $x_1, \dots, x_n$ in S and functions $f_1, \dots, f_n$ in W such that $f_i(x_j) = \delta_{ij}$.

Solution: I'm not sure using the double dual is really the easiest way to prove this. It can be done rather easily directly by induction on n (in fact see question 121704 on math.stackexchange.com). However, since H&K clearly want this done with the double dual. At first glance you might try to think of W as a dual on S and W^* as the double dual somehow. But that doesn't work since S is just a set. Instead I think you have to consider the double dual of W, W^{**} to make it work. I came up with the following solution.

Let $s \in S$. We first show that the function

$$\begin{aligned} \phi_s : &W \to F \\ &w \mapsto w(s) \end{aligned}$$

is a linear functional on W (in other words for each s, we have $\phi_s \in W^*$).

Let $w_1, w_2 \in W$, $c \in F$. Then $\phi_s(cw_1 + w_2) = (cw_1 + w_2)(s)$ which by definition equals $cw_1(s) + w_2(s)$ which equals $c\phi_s(w_1) + \phi_s(w_2)$. Thus ϕ_s is a linear functional on W.

Suppose $\phi_s(w) = 0$ for all $s \in S$, $w \in W$. Then $w(s) = 0\ \forall\ s \in S$, $w \in W$, which implies $\dim(W) = 0$. So as long as $n > 0$, $\exists\ s_1 \in S$ such that $\phi_{s_1}(w) \neq 0$ for some $w \in W$. Equivalently there is an $s_1 \in S$ and a $w_1 \in W$ such that $w_1(s_1) \neq 0$. This means $\phi_{s_1} \neq 0$ as elements of W^*. It follows that $\langle \phi_{s_1} \rangle$, the subspace of W^* generated by ϕ_{s_1}, has dimension one. By scaling if necessary, we can further assume $w_1(s_1) = 1$.

Now suppose $\forall\ s \in S$ that we have $\phi_s \in \langle \phi_{s_1} \rangle$, the subspace of W^* generated by ϕ_{s_1}. Then for each $s \in S$ there is a $c(s) \in F$ such that $\phi_s = c(s)\phi_{s_1}$ in W^*. Then for each $s \in S$, $w(s) = c(s)w(s_1)$ for all $w \in W$. In particular $w_1(s) = c(s)$ (recall $w_1(s_1) = 1$). Let $w \in W$. Let $b = w(s_1)$. Then $w(s) = c(s)w(s_1) = bw_1(s)\ \forall\ s \in S$. Notice that b depends on w but does not depend on s. Thus $w = bw_1$ as functions on S where $b \in F$ is a fixed constant. Thus $w \in \langle w_1 \rangle$, the subspace of W generated by w_1. Since w was arbitrary, it follows that $\dim(W) = 1$. Thus as long as $\dim(W) \geq 2$ we can find $w_2 \in W$ and $s_2 \in S$ such that $\langle w_1, w_2 \rangle$ (the subspace of W generated by w_1, w_2) and $\langle \phi_{s_1}, \phi_{s_2} \rangle$ (the subspace of W^* generated by $\{\phi_{s_1}, \phi_{s_2}\}$) both have dimension two. Let $W_0 = \langle w_1, w_2 \rangle$. Then we've shown that $\{\phi_{s_1}, \phi_{s_2}\}$ is a basis for W_0^*. Therefore there's a dual basis $\{F_1, F_2\} \subseteq W_0^{**}$; so that $F_i(\phi_{s_j}) = \delta_{ij}$, $i, j \in \{1, 2\}$. By Theorem 17, $\exists$ corresponding $w_1, w_2 \in W$ so that $F_i = L_{w_i}$ (in the notation of Theorem 17). Therefore, $\delta_{ij} = F_i(\phi_{s_j}) = L_{w_i}(\phi_{s_j}) = \phi_{s_j}(w_i) = w_i(s_j)$, for $i, j \in \{1, 2\}$.

Now suppose $\forall\ s \in S$ that we have $\phi_s \in \langle \phi_{s_1}, \phi_{s_2} \rangle \subseteq W^*$. Then $\forall\ s \in S$, there are constants $c_1(s), c_2(s) \in F$ and we have $w(s) = c_1(s)w(s_1) + c_2(s)w(s_2)$ for all $w \in W$. Similar to the argument in the previous paragraph, this implies $\dim(W) \leq 2$ (for $w \in W$ let $b_1 = w(s_1)$ and $b_2 = w(s_2)$ and argue as before). Therefore, as long as $\dim(W) \geq 3$ we can find s_3 so that $\langle \phi_{s_1}, \phi_{s_2}, \phi_{s_3} \rangle \subseteq W^*$, the subspace of W^* generated by $\phi_{s_1}, \phi_{s_2}, \phi_{s_3}$, has dimension three. And as before we can find $w_3 \in W$ such that $w_i(s_j) = \delta_{ij}$, for $i, j \in \{1, 2, 3\}$.

Continuing in this way we can find n elements $s_1, \ldots, s_n \in S$ such that $\phi_{s_1}, \ldots, \phi_{s_n}$ are linearly independent in W^* and corresponding elements $w_1, \ldots, w_n \in W$ such that $w_i(s_j) = \delta_{ij}$. Let $f_i = w_i$ and we are done.

Section 3.7: The Transpose of a Linear Transformation

Exercise 1: Let F be a field and let f be the linear functional on F^2 defined by $f(x_1, x_2) = ax_1 + bx_2$. For each of the following linear operators T, let $g = f^t f$, and find $g(x_1, x_2)$.

(a) $T(x_1, x_2) = (x_1, 0)$;

(b) $T(x_1, x_2) = (-x_2, x_1)$;

(c) $T(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$.

Solution:

(a) $g(x_1, x_2) = T^t f(x_1, x_2) = f(T(x_1, x_2)) = f(x_1, 0) = ax_1$.

(b) $g(x_1, x_2) = T^t f(x_1, x_2) = f(T(x_1, x_2)) = f(-x_2, x_1) = -ax_s + bx_1$.

(c) $g(x_1, x_2) = T^t f(x_1, x_2) = f(T(x_1, x_2)) = f(x_1 - x_2, x_1 + x_2) = a(x_1 - x_2) + b(x_1 + x_2) = (a + b)x_1 + (b - a)x_2$.

Exercise 2: Let V be the vector space of all polynomial functions over the field of real numbers. Let a and b be fixed real numbers and let f be the linear functional on V defined by

$$f(p) = \int_a^b p(x)dx.$$

If D is the differentiation operator on V, what is $D^t f$?

Solution: Let $p(x) = c_0 + c_1 x + \cdots + c_n x^n$. Then

$$\begin{aligned} D^t f(p) =& f(D(p)) \\ =& f(c_1 + 2c_2 x + 3c_3 x^2 + \cdots + nc_n x^{n-1}) \\ =& c_1 + c_2 x^2 + \cdots c_n x^n \mid_0^b \\ =& p(b) - p(a) \end{aligned}$$

Exercise 3: Let V be the space of all $n \times n$ matrices over a field F and let B be a fixed $n \times n$ matrix. If T is the linear operator on V defined by $T(A) = AB - BA$, and if f is the trace function, what is $T^t f$?

Solution: By exercise 3 in section 3.5, we know $\mathrm{trace}(AB) = \mathrm{trace}(BA)$. Thus $T^t f(A) = f(T(A)) = \mathrm{trace}(AB - BA) = \mathrm{trace}(AB) - \mathrm{trace}(BA) = 0$.

Exercise 4: Let V be a finite-dimensional vector space over the field F and let T be a linear operator on V. Let c be a scalar and suppose there is a non-zero vector α in V such that $T\alpha = c\alpha$. Prove that there is a non-zero linear functional f on V such that $T^t f = cf$.

Solution: Consider the operator $U = T - cI$. Then $U(\alpha) = 0$ so $\mathrm{rank}(U) < n$. Therefore $\mathrm{rank}(U^t) < n$ as an operator on V^*. It follows that there's a $f \in V^*$ such that $U^t(f) = 0$. Now $U^t = T^t - cI$, thus $T^t(f) = cf$.

Exercise 5: Let A be an $m \times n$ matrix with *real* entries. Prove that $A = 0$ if and only if $\mathrm{trace}(A^t A) = 0$.

Solution: Suppose A is the $m \times n$ matrix with entries a_{ij} and B is the $n \times k$ matrix with entries b_{ij}. Then the i, j entry of AB is

$$\sum_{k=1}^{n} a_{ik} b_{kj}.$$

Substituting A^t for A and A for B we get the i, j entry of $A^t A$ is (note that $A^t A$ has dimension $n \times n$)

$$\sum_{k=1}^{m} a_{ki} a_{kj}.$$

Thus the diagonal entries of $A^t A$ are

$$\sum_{k=1}^{m} a_{ki} a_{ki}$$

for $i = 1, \dots, n$. Thus the trace is

$$\sum_{i=1}^{n} \sum_{k=1}^{m} a_{ki}^2.$$

If all $a_{ij} \in \mathbb{R}$ then this sum is zero if and only if each $a_{ij} = 0$ because every one of them appears in this sum.

Exercise 6: Let n be a positive integer and let V be the space of all polynomial functions over the field of real numbers which have degree at most n, i.e., functions of the form

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

Let D be the differentiation operator on V. Find a basis for the null space of the transpose operator D^t.

Solution: The null space of D^t consists of all linear functionals $g : V \to \mathbb{R}$ such that $D^t g = 0$, or equivalently $g \circ D(f) = 0$ $\forall$ $f \in V$. Now $g(a_0 + a_1 x + \cdots + a_n x^n) = c_0 a_0 + c_1 a_1 + \cdots + c_n a_n$ for some consants $c_0, \dots, c_n \in \mathbb{R}$. So $g \circ D(a_0 + a_1 x + \cdots + a_n x^n) = g(a_1 + 2a_2 x + \cdots + n a_n x^{n-1}) = \sum_{i=0}^{n-1} (i+1) c_i a_{i+1}$.

This sum $\sum_{i=0}^{n-1} (i+1) c_i a_{i+1}$ equals zero for all vectors $(a_0, a_1, \dots, a_n) \in \mathbb{R}^n$ if and only if $c_0 = c_1 = \cdots = c_{n-1} = 0$. While c_n can be anything. Therefore the null space has dimension one and a basis is given by taking $c_n = 1$, which gives the function $g : \sum a_i x^i \mapsto a_n$, the projection onto the x^n coordinate.

Exercise 7: Let V be a finite-dimensional vector space over the field F. Show that $T \to T^t$ is an isomorphism of $L(V, V)$ onto $L(V^*, V^*)$.

Solution: Choose a basis $\mathcal{B}$ for V. This gives an isomorphism $L(V, V) \to M_n$ the space of $n \times n$ matrices. And the dual basis $\mathcal{B}'$ gives an isomorphism $L(V^*, V^*) \to M_n$. Now we know by Theorem 23 that the following diagram of functions commutes.

This means if we start at $L(V,V)$ and follow two functinos to the M_n in the bottom left, it doesn't matter which way around the diagram we go, we end up at the same place.

$$\begin{array}{rcl} L(V,V) & \longrightarrow & M_n \\ \text{transpose} \downarrow & & \downarrow \text{transpose} \\ L(V^*,V^*) & \longrightarrow & M_n \end{array}$$

Both horizontal arrows are isomorphisms (by Theorem 12, page 88). Now clearly transpose on matrices is a one-to-one and onto function from the set of matrices to itself. Also $(rA)^t = rA^t$ and $(A+B)^t = A^t + B^t$ for any two $n \times n$ matrices A and B. Thus transpose is also a linear transformation. Thus transpose is an isomorphism. Therefore three of the arrows in this diagram are isomorphisms and it follows that the fourth arrow must also be an isomorphism.

Exercise 8: Let V be the vector space of $n \times n$ matrices over the field F.

(a) If B is a fixed $n \times n$ matrix, define a function f_B on V by $F_B(A) = \text{trace}(B^tA)$. Show that F_B is a linear functional on V.

(b) Show that every linear functional on V is of the above form, i.e., is f_B for some B.

(c) Show that $B \to f_B$ is an isomorphism of V onto V^*.

Solution:

(a) This follows from the fact that the trace function is a linear functional and left multiplication by a matrix is a linear transformation from V to V. In other words $F_B(cA_1+A_2) = \text{trace}(B^t(cA_1+A_2)) = \text{trace}(cB^tA_1+B^tA_2) = c\cdot\text{trace}(B^tA_1)+\text{trace}(B^tA_2) = c \cdot F_B(A_1) + F_B(A_2)$.

(b) Let $f : V \to F$ be a linear functional. Let $A = (a_{ij}) \in V$. Then

$$f(A) = \sum_{i,j=1}^{n} c_{ij}a_{ij} \tag{23}$$

for some fixed constants $c_{ij} \in F$.

Now let $B = (b_{ij}) \in V$ be any matrix. Then the i, j element of B^tA is $\sum_{k=1}^{n} b_{ki}a_{kj}$. Thus

$$\text{trace}(B^tA) = \sum_{i=1}^{n}\sum_{k=1}^{n} b_{ki}a_{ki}. \tag{24}$$

Comparing (23) and (24) we see each a_{ij} apperas exactly once in each sum. So setting $b_{ki} = c_{ki}$ for all $i,k = 1,\dots,n$ we get the appropriate matrix B such that $\text{trace}(B^tA) = f$.

(c) Let F be the function $F : V \to V^*$ such that $F(B) = f_B$. Part (a) shows this function is into V^*. Part (b) shows it is onto V^*. We must show it is linear and one-to-one. Let $r \in F$ and $B_1, B_2 \in V$. Then $(rB_1 + B_2)^tA = (rB_1^t + B_2^t)A = rB_1^tA + B_2^tA$. We know the trace function itself is linear. Thus F is linear. Now suppose $\text{trace}(B^tA) = 0 \ \forall\ A \in V$. Fix $i, j \in \{1, 2, \cdots, n\}$. Let A be the matrix with a one in the i, j position and zeros elsewhere. Then by the proof of part (b) we know that $\text{trace}(B^tA) = b_{ij}$. So if $F(A) = 0$ then $b_{ij} = 0$. Thus B must be the zero matrix. Thus F is one-to-one. It follows that F is an isomorphism.

Chapter 4: Polynomials

Section 4.2: The Algebra of Polynomials

Exercise 1: Let F be a subfield of the complex numbers and let A be the following 2×2 matrix over F

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

For each of the following polynomials f over F, compute $f(A)$.

(a) $f = x^2 - x + 2$;

(b) $f = x^3 - 1$;

(c) $f = x^2 - 5x + 7$;

Solution:

(a)

$$\begin{aligned} A^2 &= \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 5 \\ -5 & 8 \end{bmatrix} \end{aligned}$$

Therefore

$$\begin{aligned} f(A) &= A^2 - A + 2 \\ &= \begin{bmatrix} 3 & 5 \\ -5 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 \\ -4 & 7 \end{bmatrix} \end{aligned}$$

(b)

$$\begin{aligned} A^2 &= \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 5 \\ -5 & 8 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 18 \\ -18 & 19 \end{bmatrix} \end{aligned}$$

Therefore

$$f(A) = A^3 - 1$$

$$= \begin{bmatrix} 1 & 18 \\ -18 & 19 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 18 \\ -18 & 18 \end{bmatrix}$$

(c)

$$f(A) = A^2 - 5A + 7$$

$$= \begin{bmatrix} 3 & 5 \\ -5 & 8 \end{bmatrix} - 5\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Exercise 2: Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T(x_1, x_2, x_3) = (x_1, x_3, -2x_2 - x_3).$$

Let f be the polynomial over $\mathbb{R}$ defined by $f = -x^3 + 2$. Find $f(T)$.

Solution: The matrix of T with respect to the standard basis is

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

Thus

$$A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 2 & -1 \end{bmatrix}$$

So

$$A^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -1 \end{bmatrix}$$

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

Thus

$$-A^3 + 2 = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

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

This corresponds to the transformation

$$f(T)(x_1, x_2, x_3) = (-x_1, x_3, -x_2 - x_3).$$

Exercise 3: Let A be an $n \times n$ diagonal matrix over the field F, i.e., a matrix satisfying $A_{ij} = 0$ for $i \neq j$. Let f be the polynomial over F defined by

$$f = (x - A_{11}) \cdots (x - A_{nn}).$$

What is the matrix $f(A)$?

Solution: The product of n diagonal matrices is again a diagonal matrix, where the i, i element is the product of the entries in the i, i position of the n individual matrices. For each i, the i, i terms of $A - A_{ii}$ is zero. Therefore each diagonal entry of $(A - A_{11}) \cdots (A - A_{nn})$ is a product of numbers one of which is zero. Therefore $(A - A_{11}) \cdots (A - A_{nn})$ is the zero matrix.

Exercise 4: If f and g are independent polynomials over the field F and h is a non-zero polynomial over F, show that fh and gh are independent.

Solution: Suppose there are scalars $r, s \in F$ such that $rfh + sgh =$. Then $rfh = -sgh$ and so by Corollary 2 on page 121, it follows that $rf = -sg$ so that $rf + sg = 0$. Since f and g are independent, it follows that $r = s = 0$. Thus fh and gh are independent.

Exercise 5: If F is a field, show that the product of two non-zero elements of F^∞ is non-zero.

Solution: Let $f = (f_0, f_1, f_2, \dots)$ and $g = (g_0, g_1, g_2, \dots)$ be two elements of F^∞. Let n be the index of the first non-zero f_i and let m be the index of the first non-zero g_i (could be $n = m = 0$). Then what is the $n + m$ coordinate of the product? It is

$$(fg)_{n+m} = \sum_{i=0}^{n+m} f_i g_{n+m-i} = \sum_{i=n}^{n+m} f_i g_{n+m-i}$$

and if $i > n$ then $g_{n+m-i} = 0$ thus

$$\sum_{i=n}^{n+m} f_i g_{n+m-i} = \sum_{i=n}^{n} f_i g_{n+m-i} = f_n g_m.$$

Now we've assumed f_n and g_m are non-zero thus $f_n g_m \neq 0$ and thus $fg \neq 0$ in F^∞.

Exercise 6: Let S be a set of non-zero polynomials over a field F. If no two elements of S have the same degree, show that S is an independent set in $F[x]$.

Solution: Let $f_1, \dots, f_n \in S$. Suppose $\deg(f_i) = d$ and $\deg(f_j) < d \ \forall\ i \neq j$. We can do this because by assumption all polynomials in the set $\{f_1, \dots, f_n\}$ have different degrees, so one of them must have the largest. Suppose the d-th coefficient of f_i is $r \neq 0$. Then any linear combination $a_1 f_1 + \cdots a_n f_n$ has d-th coefficient equal to ra_i. Thus if this linear combination is zero then necessarily $a_i = 0$. We can now apply the same argument to the f_j with the second largest degree to show its coefficient in the linear combination is zero. And then to the third largest, etc. Until we have eventually shown all coefficients in the linear combination are zero. It follows that $\{f_1, \dots, f_n\}$ is a linearly independent subset of S and since it was an arbitrary finite subset of S it follows that S is a linearly independent set.

Exercise 7: If a and b are elements of a field F and $a \neq 0$, show that the polynomials 1, $ax + b$, $(ax + b)^2$, $(ax + b)^3, \dots$ form a basis of $F[x]$.

Solution: Let $S = \{1, ax + b, (ax + b)^2, (ax + b)^3, \dots\}$. And let $\langle S \rangle$ be the subspace spanned by S. By the previous exercise we know S is a linearly independent set. We must just show S spans the space of all polynomials. Since $1 \in S$ and $ax + b \in S$ it follows that $b \cdot 1 + \frac{1}{a}(a + bx) \in \langle S \rangle$. Thus $x \in \langle S \rangle$. Now we can subtract a multiple of 1 and a multiple of x from $(a + bx)^2$ to get $a^2x^2 \in \langle S \rangle$. Thus $\frac{1}{a^2} \cdot a^2x^2 \in \langle S \rangle$. Thus $x^2 \in S$. Continuing in this way we can show that $x^n \in \langle S \rangle$ for all n. Since $\{1, x, x^2, \dots\}$ span the space of all polynomials, it follows that S spans the space of all polynomials.

Exercise 8: If F is a field and h is a polynomial over F of degree ≥ 1, show that the mapping $f \to f(h)$ is a one-one linear transformation of $F[x]$ into $F[x]$. Show that this transformation is an isomorphism of $F[x]$ onto $F[x]$ if and only if $\deg h = 1$.

Solution: Let $G : F[x] \to F[x]$ be the function $G(f) = f(h)$. Clearly G is a well-defined function from $F[x]$ to $F[x]$. By definition $G(f+g) = (f+g)(h) = f(h) + g(h) = G(f) + G(g)$ and for $r \in F$, $G(rf) = (rf)(h) = r \cdot f(h) = rG(F)$. Thus G is a linear transformation. Suppose $\deg h > 1$. Then the coefficient of x in $f(h)$ is zero. Thus if $\deg h > 1$ then G is not onto. Now suppose $\deg h = 0$. Then $f(h)$ is a scalar for all f. Thus G is not onto. Now suppose $\deg h = 1$, so that $h(x) = a + bx$. Let $h' = \frac{1}{b}x - \frac{a}{b}$ and let G' be the corresponding function on $F[x]$, so $G' : F[x] \to F[x]$ is given by $G(f) = f(\frac{1}{b}x - \frac{a}{b})$. Then $G \circ G'$ and $G' \circ G$ both give the identify function on $F[x]$. Thus G is an isomorphism.

Exercise 9: Let F be a subfield of the complex numbers and let T, D be the transformations on $F[x]$ defined by

$$T\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=0}^{n} \frac{c_i}{1+i} x^{i+1}$$

and

$$D\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=1}^{n} i c_i x^{i-1}.$$

(a) Show that T is a non-singular linear operator on $F[x]$. Show also that T is not invertible.

(b) Show that D is a linear operator on $F[x]$ and find its null space.

(c) Show that $DT = I$, and $TD \neq I$.

(d) Show that $T[(Tf)g] = (Tf)(Tg) - T[f(Tg)]$ for all f, g in $F[x]$.

(e) State and prove a rule for D similar to the one given for T in (d)

(f) Suppose V is a non-zero subspace of $F[x]$ such that Tf belongs to V for each f in V. Show that V is not finite-dimensional.

(g) Suppose V is a finite-dimensional subspace of $F[x]$. Prove there is an integer $m \geq 0$ such that $D^m f = 0$ for each f in V.

Solution:

(a) Clearly T is a function from $F[x]$ to $F[x]$. We must show T is linear.

$$T\left(\sum_{i=0}^{n} c_i x^i + \sum_{i=0}^{n} c_i' x^i\right) = T\left(\sum_{i=0}^{n} (c_i + c_i') x^i\right)$$

$$= \sum_{i=0}^{n} \frac{c_i + c_i'}{i+1} x^{i+1}$$

$$= \sum_{i=0}^{n} \left(\frac{c_i}{1+i} x^{i+1} + \frac{c_i'}{1+i} x^{i+1}\right)$$

$$= T\left(\sum_{i=0}^{n} c_i x^i\right) + T\left(\sum_{i=0}^{n} c_i' x^i\right).$$

and

$$T\left(r \sum_{i=0}^{n} c_i x^i\right) = T\left(\sum_{i=0}^{n} r c_i x^i\right) = \sum_{i=0}^{n} \frac{r c_i}{i+1} x^{i+1} = r \sum_{i=0}^{n} \frac{c_i}{i+1} x^{i+1} = r \cdot T\left(\sum_{i=0}^{n} c_i x^i\right).$$

Thus T is linear.

Since F has characterisitc zero we can find $a, b \in F$, such that $a \neq b$. Consider a and b as constant polynomials in F. Then $T(a) = T(b) = 0$. Thus T is not one-to-one. Thus T is not invertible.

(b) Clearly D is a function from $F[x]$ to $F[x]$. We must show D is linear.

$$D\left(\sum_{i=0}^{n} c_i x^i + \sum_{i=0}^{n} c_i' x^i\right) = D\left(\sum_{i=0}^{n} (c_i + c_i')x^i\right)$$

$$= \sum_{i=1}^{n} i(c_i + c_i')x^{i-1}$$

$$= \sum_{i=1}^{n} (ic_i x^{i-1} + ic_i' x^{i-1})$$

$$= D\left(\sum_{i=0}^{n} c_i x^i\right) + D\left(\sum_{i=0}^{n} c_i' x^i\right).$$

and

$$D\left(r\sum_{i=0}^{n} c_i x^i\right) = D\left(\sum_{i=0}^{n} rc_i x^i\right) = \sum_{i=1}^{n} rc_i x^{i-1} = r\sum_{i=1}^{n} c_i x^{i-1} = r \cdot D\left(\sum_{i=0}^{n} c_i x^i\right).$$

Thus D is linear.

Suppose $f(x) = \sum_{i=0}^{n} c_i x^i$ is in the null space of D. Then $D(f) = \sum_{i=1}^{n} ic_i x^{i-1} = 0$. A polynomial is zero if and only if every coefficient is zero. Thus it must be that $0 = c_1 = c_2 = c_3 = \cdots$. So it must be that $f(x) = c_0$ a constant polynomial. Thus the null space of D contains the constant polynomials. Since $D(f) = 0$ for all constant polynomials, the null space of D consists of exatly the constant polynomials.

(c)

$$D\left(T\left(\sum_{i=0}^{n} c_i x^i\right)\right)$$

$$= D\left(\sum_{i=0}^{n} \frac{c_i}{1+i} x^{i+1}\right).$$

The first non-zero term of this sum is the linear term $c_0 x$. Thus when we apply D the sum still starts at zero:

$$= \sum_{i=0}^{n} (i+1)\frac{c_i}{1+i} x^{i+1-1}$$

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

Thus

$$D\left(T\left(\sum_{i=0}^{n} c_i x^i\right)\right) = \sum_{i=0}^{n} c_i x^i.$$

Thus $DT = I$.

Let $f(x) = 1$. Then $TD(f) = T(D(f)) = T(0) = 0$. Thus $TD(1) \neq 1$. Thus $TD \neq I$.

(d) This follows rather easily from part (e). And likewise (e) follows rather easily from (d). Thus one can derive (d) straight from the definition of T and then derive (e) from it, or one can derive (e) straight from the definition of D and then derive (d) from it. I've chosen to do the latter.

In part (e) below the product formula is proven straight from the definition. So we will use it here to prove this part. In particular, we apply the product formula from part (e) to $(Tf)(Tg)$

$$D[(Tf)(Tg)] = (Tg)D(Tf) + (Tf)D(Tg).$$

By part (c) $DT = I$ so this is equivalent to

$$D[(Tf)(Tg)] = f(Tg) + (Tf)g.$$

Thus

$$D[(Tf)(Tg)] - f(Tg) = (Tf)g.$$

Now apply T to both sides

$$T\Big(D[(Tf)(Tg)] - f(Tg)\Big) = T((Tf)g).$$

Since T is a linear transformation this is equivalent to

$$T\Big(D[(Tf)(Tg)]\Big) - T[f(Tg)] = T((Tf)g).$$

We showed in part (c) that $TD \neq I$, however if f has constant term equal to zero then in fact $T(D(f))$ does equal f. Now Tf and Tg have constant term equal to zero, so $(Tf)(Tg)$ has constant term zero, thus

$$T\Big(D[(Tf)(Tg)]\Big) = (Tf)(Tg).$$

Thus

$$(Tf)(Tg) - T[f(Tg)] = T((Tf)g).$$

(e) I believe they are after the product formula here:

$$D(fg) = fD(g) + gD(f). \tag{25}$$

We prove this by brute force appealing just to the definition and to the product formula for polynomials. Let $f(x) = \sum_{i=0}^{n} c_i x^i$ and $g(x) = \sum_{i=0}^{m} d_i x^i$. Then using the product formula (4-8) on page 121 we have

$$D(fg) = D\left(\sum_{i=0}^{n} c_i x^i \sum_{i=0}^{m} d_i x^i\right) = D\left(\sum_{i=0}^{n+m}\left(\sum_{j=0}^{i} c_j d_{i-j}\right) x^i\right)$$

And using the linearity of the differentiation operator D this equals

$$\sum_{i=0}^{n+m} i\left(\sum_{j=0}^{i} c_j d_{i-j}\right) x^{i-1}. \tag{26}$$

Now we write down the sum for the right hand side of (25):

$$\begin{aligned} & fD(g) + gD(f) \\ &= \sum_{i=0}^{n} c_i x^i \sum_{i=1}^{m} i d_i x^{i-1} + \sum_{i=1}^{n} i c_i x^{i-1} \sum_{i=0}^{m} d_i x^i. \end{aligned} \tag{27}$$

Consider

$$x \cdot \sum_{i=0}^{n} c_i x^i \sum_{i=1}^{m} i d_i x^{i-1}.$$

This equals

$$\sum_{i=0}^{n} c_i x^i \sum_{i=1}^{m} i d_i x^i$$

and since $0d_0 = 0$ we can write this as

$$\sum_{i=0}^{n} c_i x^i \sum_{i=0}^{m} i d_i x^i.$$

It's straightforward to apply (4-8) page 121 to this product. In (4-8) we let $f_i = c_i$ and $g_i = id_i$ and it equals

$$\sum_{i=0}^{m+n}\left(\sum_{j=0}^{i}(i-j)c_j d_{i-j}\right)x^i.$$

The constant terms is zero thus we can write it as

$$\sum_{i=1}^{m+n}\left(\sum_{j=0}^{i}(i-j)c_j d_{i-j}\right)x^i.$$

And thus the sum

$$\sum_{i=0}^{n} c_i x^i \sum_{i=1}^{m} i d_i x^{i-1}$$

equals

$$\sum_{i=1}^{m+n}\left(\sum_{j=0}^{i}(i-j)c_j d_{i-j}\right)x^{i-1}.$$

Similary the second sum is

$$\sum_{i=1}^{n+m}\left(\sum_{j=0}^{i} j c_j d_{i-j}\right)x^i.$$

Thus (27) does equal (26).

(f) Suppose V is finite dimensional. Let $\{b_1, \cdots, b_n\}$ be a basis for V. Let $d = \max_{i=1,\cdots n}\deg(b_i)$. It follows from Theorem 1(v) and induction that the degree of a linear combination of polynomials is no larger than the max of the degress of the individual polynomials involved in the linear combination. Thus no element of V has degree greater than d. Now let $f \in V$ be any non-zero element. Let $d' = \deg(f)$. Then Tf has degree $d' + 1$, $T^2 f$ has degree $d' + 2$, etc. Thus for some n, $\deg(T^n f) > d$. If $T^n f \in V$ then this is a contradiction. Thus if $T^n f \in V$ for all $f \in V$ it must be that V is not finite dimensional.

(g) Let $\{b_1, \cdots, b_n\}$ be a basis for V. Let $d = \max_{i=1,\cdots n}\deg(b_i)$. For any $f \in F[x]$, we know $\deg(Df) < \deg(f)$. Thus $D^{d+1}b_i = 0$ for all $i = 1, \ldots, n$. Since $D^{d+1}(b_i) = 0$ for all elements of the basis $\{b_1, \ldots, b_n\}$ it follows that $D^{d+1}(f) = 0$ for all $f \in V$.

Section 4.3: Lagrange Interpolation

Exercise 1: Use the Lagrange interpolation formula to find a polynomial f with real coefficients such that f has degree ≤ 3 and $f(-1) = -6$, $f(0) = 2$, $f(1) = -2$, $f(2) = 6$.

Solution: $t_0 = -1$, $t_1 = 0$, $t_2 = 1$, $t_3 = 2$. Therefore

$$\begin{aligned} P_0 &= \frac{x(x-1)(x-2)}{(-1)(-2)(-3)} = \frac{-1}{6}x(x-1)(x-2) \\ P_1 &= \frac{(x+1)(x-1)(x-2)}{(-1)(-2)} = \frac{1}{2}(x-1)(x+1)(x-2) \\ P_2 &= \frac{(x+1)x(x-2)}{(2)(1)(-1)} = \frac{-1}{2}x(x+1)(x-2) \\ P_3 &= \frac{(x+1)x(x-1)}{(3)(2)(1)} = \frac{1}{6}(x-1)(x+1). \end{aligned}$$

Thus

$$\begin{aligned} f &= f(-1)\cdot P_0 + f(0)\cdot P_1 + f(1)\cdot P_2 + f(2)\cdot P_3 \\ &= -6P_0 + 2P_1 - 2P_2 + 6P_3 \\ &= x(x-1)(x-2) + (x-1)(x+1)(x-2) + x(x+1)(x-2) + x(x-1)(x+1) \\ &= (x^3 - 3x^2 + 2x) + (x^3 - 2x^2 - x + 2) + (x^3 - x^2 - 2x) + (x^3 - x) \\ &= 4x^3 - 6x^2 - 2x + 2. \end{aligned}$$

Checking:

$$\begin{aligned} f(-1) &= -4 - 6 + 2 + 2 = -6 \\ f(0) &= 2 \\ f(1) &= 4 - 6 - 2 + 2 = -2 \\ f(2) &= 32 - 24 - 4 + 2 = 6. \end{aligned}$$

Exercise 2: Let $\alpha, \beta, \gamma, \delta$ be real numbers. We ask when it is possible to find a polynomial f over $\mathbb{R}$, of *degree not more than* 2, such that $f(-1) = \alpha$, $f(1) = \beta$, $f(3) = \gamma$ and $f(0) = \delta$. Prove that this is possible if and only if

$$3\alpha + 6\beta - \gamma - 8\delta = 0.$$

Solution: Let $t_0 = -1$, $t_1 = 1$, $t_2 = 3$. We will apply the Lagrange interpolation formula to get a quadratic satisfying $f(t_0) = \alpha$, $f(t_1) = \beta$, $f(t_2) = \gamma$. Then we will figure out what condition on $\alpha, \beta, \gamma, \delta$ will guarantee that it also satisfies $f(0) = \delta$.

$$\begin{aligned} P_0 &= \frac{(x-1)(x-3)}{(-2)(-4)} = \frac{1}{8}(x-1)(x-3) \\ P_1 &= \frac{(x+1)(x-3)}{(2)(-2)} = \frac{-1}{4}(x+1)(x-3) \\ P_2 &= \frac{(x+1)(x-1)}{(4)(2)} = \frac{1}{8}(x+1)(x-1) \end{aligned}$$

Therefore

$$\begin{aligned} f &= \frac{\alpha}{8}(x-1)(x-3) - \frac{\beta}{4}(x+1)(x-3) + \frac{\delta}{8}(x+1)(x-1) \\ &= \tfrac{1}{8}(\alpha x^2 - 4\alpha x + 3\alpha - 2\beta x^2 + 4\beta x + 6\beta + \gamma x^2 - \gamma). \end{aligned}$$

Now $f(0) = \gamma$ implies

$$\frac{1}{8}(3\alpha + 6\beta - \gamma) = \delta.$$

Simplifying gives

$$3\alpha + 6\beta - \gamma - 8\delta = 0. \tag{28}$$

Thus if (28) is satisfied then the four values of f are as required. Since three points determine a quadratic, there cannot be any quadratic other than f that goes through $(-1,\alpha)$, $(1,\beta)$, $(3,\delta)$. Thus this condition is not only sufficient but it is necessary.

Exercise 3: Let F be the field of real numbers,

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$p = (x-2)(x-3)(x-1).$$

(a) Show that $p(A) = 0$.

(b) Let P_1, P_2, P_3 be the Lagrange polynomials for $t_1 = 2$, $t_2 = 3$, $t_2 = 1$. Compute $E_i = P_i(A)$, $i = 1, 2, 3$.

(c) Show that $E_1 + E_2 + E_3 = I$, $E_iE_j = 0$ if $i \neq j$, $E_i^2 = E_i$.

(d) Show that $A = 2E_1 + 3E_2 + E_3$.

Solution: (a)

$$(A-2)(A-3)(A-1)$$
$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(b) $t_1 = 2$, $t_2 = 3$, $t_3 = 1$.

$$P_1 = -(x-3)(x-1)$$
$$P_2 = \tfrac{1}{2}(x-2)(x-1)$$
$$P_3 = \tfrac{1}{2}(x-2)(x-3)$$

Thus

$$E_1 = P_1(A) = -(A-3I)(A-I) = -\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$E_2 = P_2(A) = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$E_3 = P_3(A) = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(c) All three of these facts are basically obvious by visual inspection of the matrices in part (b). $E_1 + E_2 + E_3 = I$ is obvious by inspection. Likewise it is evident by inspection that $E_iE_j = 0$ if $i \neq j$. Lastly it is obvious that $E_i^2 = E_i$. I'm not sure what there is to prove here.

(d)

$$\begin{aligned} &2E_1 + 3E_2 + E_3 \\ &= 2\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Exercise 4: Let $p = (x-2)(x-3)(x-1)$ and let T be any linear operator on $\mathbb{R}^4$ such that $p(T) = 0$. Let P_1, P_2, P_3 be the Lagrange polynomials of Exercise 3, and let $E_i = P_i(T)$, $i = 1, 2, 3$. Prove that

$$E_1 + E_2 + E_3 = I, \quad E_iE_j = 0 \quad \text{if} \quad i \neq j,$$

$$E_i^2 = E_i, \quad \text{and} \quad T = 2E_1 + 3E_2 + E_3.$$

Solution: Recall

$$\begin{aligned} P_1 &= -x^2 + 4x - 3 \\ P_2 &= \tfrac{1}{2}x^2 - \tfrac{3}{2}x + 1 \\ P_2 &= \tfrac{1}{2}x^2 - \tfrac{5}{2}x + 3 \end{aligned}$$

By definition $P(T) + Q(T) = (P+Q)(T)$ and $P(T)Q(T) = (PQ)(T)$. Now as polynomials $P_1 + P_2 + P_3 = 1$. Thus $E_1 + E_2 + E_3 = P_1(T) + P_2(T) + P_3(T) = (P_1 + P_2 + P_3)(T) = I$.

Notice that P divides P_iP_j whenever $i \neq j$. Thus $P_iP_j = PQ$ for some Q (as polynomials). Thus $E_iE_j = P_i(T)P_j(T) = (P_iP_j)(T) = P(T)Q(T) = 0 \cdot Q(T) = 0$. Thus $E_iE_j(T) = 0$.

We prove the next part in general. Let

$$\begin{aligned} f_1 &= x - a \\ f_2 &= x - b \\ f_3 &= x - c \end{aligned}$$

Thus

$$\begin{aligned} P_1 &= \frac{f_2f_3}{(a-b)(a-c)} \\ P_2 &= \frac{f_1f_3}{(b-a)(b-c)} \\ P_3 &= \frac{f_1f_2}{(c-a)(c-b)} \end{aligned}$$

Let $d = 2c - b - a$. Then it follows by simply multiplying it out that

$$\frac{df_3 + f_3^2}{(c-a)(c-b)} = \frac{f_1 f_2}{(c-a)(c-b)} - 1.$$

Which is equivalent to

$$\frac{df_3 + f_3^2}{(c-a)(c-b)} = P_3 - 1. \tag{29}$$

This equation is true as polynomials. We now evaluate things at T.

$$f_1(T)f_2(T)f_3(T) = 0$$

$$\Longrightarrow$$

$$f_1(T)f_2(T)f_3^2(T) = 0$$

$$\Longrightarrow$$

$$\frac{f_1(T)f_2(T)}{(c-a)(c-b)} \cdot \frac{f_3^2(T)}{(c-a)(c-b)} = 0.$$

Since $f_1(T)f_2(T)f_3(T) = 0$, this implies

$$\frac{f_1(T)f_2(T)}{(c-a)(c-b)} \cdot \frac{df_3(T) + f_3^2(T)}{(c-a)(c-b)} = 0.$$

Equivalently

$$P_3(T) \cdot \frac{df_3(T) + f_3^2(T)}{(c-a)(c-b)} = 0.$$

By (29) this implies

$$P_3(T)(P_3(T) - 1) = 0.$$

Thus

$$P_3^2(T) = P_3(T).$$

Thus $E_3^2 = E_3$. Since a, b, c were general, the same follows for E_1 and E_2.

It remains to show $T = 2E_1 + 3E_2 + E_3$. We first note that as polynomials

$$2P_1 + 3P_2 + P_3$$

$$= (-2x^2 + 8x - 6) + (\tfrac{3}{2}x^2 - \tfrac{9}{2}x + 3) + (\tfrac{1}{2}x^2 - \tfrac{5}{2}x + 3)$$

$$= x.$$

Plugging in T we get

$$2P_1(T) + 3P_2(T) + P_3(T) = T.$$

Thus

$$2E_1 + 3E_2 + E_3 = T.$$

.

Exercise 5: Let n be a positive integer and F a field. Suppose T is an $n \times n$ matrix over F and P is an invertible $n \times n$ matrix over F. If f is any polynomial over F, prove that

$$f(P^{-1}TP) = P^{-1}f(T)P.$$

Solution: First note that $(P^{-1}xP)^n = P^{-1}x^n(T)P$. This is obvious by inspection, it follows basically from the fact that multiplication is associative and $P^{-1}P = I$.

The general result now follows

$$\begin{aligned}&P^{-1}f(T)P\\&= P^{-1}(a_0 + a_1T + a_2T^2 + \cdots + a_nT^n)P\\&= P^{-1}a_0P + P^{-1}a_1TP + P^{-1}a_2T^2P + \cdots + P^{-1}a_nT^nP\\&= a_0 + a_1(P^{-1}TP) + a_2(P^{-1}TP)^2 + \cdots + a_n(P^{-1}TP)^n\\&= f(P^{-1}TP).\end{aligned}$$

Exercise 6: Let F be a field. We have considered certain special linear functionals on $F[x]$ obtained via 'evaluation at t':

$$L(f) = f(t).$$

Such functionals are not only linear but also have the property that $L(fg) = L(f)L(g)$. Prove that if L is any linear functional on $F[x]$ such that

$$L(fg) = L(f)L(g)$$

for all f and g, then either $L = 0$ or there is a t in F such that $L(f) = f(t)$ for all f.

Solution: Let L be a non-zero linear transformation. First note that $L(1) \neq 0$ since otherwise $L(f) = L(f \cdot 1) = L(f)L(1) = L(f) \cdot 0 = 0\ \forall\ f$. Next note that $L(1) = L(1 \cdot 1) = L(1)L(1) \Rightarrow L(1) = 1$. It follows that $L(a) = L(a \cdot 1) = aL(1) = a\ \forall\ a \in F$. Now let $t = L(x)$. Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$. Then

$$\begin{aligned}&L(f)\\&= L(a_0 + a_1x + \cdots + a_nx^n)\\&= L(a_0) + L(a_1)L(x) + L(a_2)L(x^2)\cdots + L(a_n)L(x^n)\\&= a_0 + a_1L(x) + a_2L(x)^2 + \cdots + a_nL(x)^n\\&= a_0 + a_1t + a_2t^2 + \cdots + a_nt^n\\&= f(t).\end{aligned}$$

Section 4.4: Polynomial Ideals

Page 129: In the statement of Theorem 5 it says "If f is a polynomial over f" which I believe should be "If f is a polynomial over F".

Page 133: The page's running title says "*Polynomial Ideas*" it should say "*Polynomial Ideals*".

Page 134: In exercise 2 (a) and (c) I think they made a calculation mistake because both problems are extremely tedious to do with only what we know so far. I believe in part (a) they really meant $2x^5 - x^3 - 3x^2 - 6x + 6$ and in part (c) they really meant $x^3 + 6x^2 + 7x + 2$. If I were teaching out of this book I would change both problems to be like this. In any case, I solved them below as stated in the book, for the sake of completeness. But the derivations are rather ugly.

Exercise 1: Let $\mathbb{Q}$ be the field of rational numbers. Determine which of the following subsets of $\mathbb{Q}[x]$ are ideals. When the set is an ideal, find its monic generator.

(a) all f of even degree;

(b) all f of degree ≥ 5;

(c) all f such that $f(0) = 0$;

(d) all f such that $f(2) = f(4) = 0$;

(e) all f in the range of the linear operator T defined by

$$T\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=0}^{n} \frac{c_i}{i+1} x^{i+1}.$$

Solution:

(a) This is not an ideal. Let $f(x)$ be any polynomial of even degree. Let $g(x) = x$. Then $\deg(fg) = \deg(f) + 1$ which is odd and is therefore the set of polynomials of even degree does not satisfy the necessary property that fg is in the set whenever f is in the set.

(b) This is not an ideal. Let $f(x) = x^5$ and $g(x) = -x^5 + x^4$. Then f and g are in the set and therefore $f + g$ must be in the set if it is an ideal. But $f(x) + g(x) = x^4$ has degree equal to four and therefore is not in the set. In other words the set is not closed with respect to addition and therefore is not even a subspace.

(c) This is an ideal. Let $I = \{f \in F[x] \mid f(0) = 0\}$. If $f(0) = 0$ and $g(0) = 0$ then $(cf + g)(0) = cf(0) + g(0) = 0$ thus I is a subspace of $F[x]$. Now suppose $f \in I$ and $g \in F[x]$. Then $(fg)(0) = f(0)g(0) = 0 \cdot g(0) = 0$. Thus $fg \in I$. Thus I is an ideal. Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$. Then $f(0) = a_0$. Thus if $f \in I$ then necessarily $a_0 = 0$. Let $g(x) = x$. Then $g \in I$ and $f(x) = a_1x + a_2x^2 + \cdots a_nx^n = x(a_1 + a_2x + \cdots a_nx^{n-1}) \in I$. Thus g generates I.

(d) This is an ideal. Let $I = \{f \in F[x] \mid f(2) = f(4) = 0\}$. If $f(2) = f(4) = 0$ and $g(2) = g(4) = 0$ then $(cf + g)(2) = cf(2) + g(2) = 0$ and $(cf + g)(4) = cf(4) + g(4) = 0$ thus I is a subspace of $F[x]$. Now suppose $f \in I$ and $g \in F[x]$. Then $(fg)(2) = f(2)g(2) = 0 \cdot g(2) = 0$ and $(fg)(4) = f(4)g(4) = 0 \cdot g(4) = 0$. Thus $fg \in I$. Thus I is an ideal. Let $g(x) = (x-2)(x-4)$. We claim g generates I. Let $f(x) \in I$. Then since $f(2) = 0$, by Corollary 1 page 128 it follows that $f(x) = (x-2)q(x)$ for some $q(x) \in F[x]$. Now since $f(4) = 2 \cdot q(4) = 0$ it follows that $q(4) = 0$ Thus there is a $p(x)$ such that $q(x) = (x-4)p(x)$. Thus $f(x) = (x-2)(x-4)p(x) = g(x)p(x)$ and therefore $f(x)$ is in the ideal generated by $g(x)$.

(e) This is an ideal. In fact it is the same ideal as in part (c). Note that $T(f)$ has no constant term, thus $T(f)(0) = 0\ \forall\ f \in F[x]$. Now let $f(x) = a_1x + a_2x^2 + \cdots + a_nx^n$. Let $g(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots + na_nx^{n-1}$. Then $f = T(g)$. Thus f is in the image of T. Thus any polynomial with zero constant term is in the image of T. Thus it is exactly the same ideal is in part (c).

Exercise 2: Find the g.c.d. of each of the following paris of polynomials

(a) $2x^5 - x^3 - 3x^2 - 6x + 4$, $x^4 + x^3 - x^2 - 2x - 2$;

(b) $3x^4 + 8x^2 - 3$, $x^3 + 2x^2 + 3x + 6$;

(c) $x^4 - 2x^3 - 2x^2 - 2x - 3$, $x^3 + 6x^2 + 7x + 1$.

Solution:

This question on its face seems to depend on what the base field is. I'm going to assume it is $\mathbb{C}$.

(a) Let $f(x) = 2x^5 - x^3 - 3x^2 - 6x + 4$ and $g(x) = x^4 + x^3 - x^2 - 2x - 2$. I have a feeling there's a mistake and they really meant $2x^5 - x^3 - 3x^2 - 6x + 6$ because then both polynomials are divisible by $x^2 - 2$. But as it is, the g.c.d. of these two polynomials equals 1 - which does not seem that easy to prove using only what we know up to this point. If we knew about unique factorization in $\mathbb{C}[x]$ we could simply factor $g(x)$ into linear factors and check none of the roots of g are roots of f. In fact the roots of $g(x)$ are $\pm\sqrt{2}$ and $\frac{-1\pm\sqrt{2}i}{2}$.

But we can't use that argument yet. Instead we are probably expected to argue using the comments on page 132 right after the proof of Theorem 7, commonly known as the "euclidean algorithm".

Since I believe there's a mistake and it should be $2x^5 - x^3 - 3x^2 - 6x + 6$, I'm going to solve it both ways.

First we solve it exactly as stated in the book.

Solution 1 (as stated in book).

Let $I = \langle f(x), g(x) \rangle$ be the ideal in $\mathbb{C}[x]$ generated by f and g. Then

$$2x^5 - x^3 - 3x^2 - 6x + 4 = (x^4 + x^3 - x^2 - 2x - 2)(2x - 2) + (3x^3 - x^2 - 6x).$$

Therefore $3x^3 - x^2 - 6x = f(x) - (2x - 2)g(x) \in I$.

$$x^4 + x^3 - x^2 - 2x - 2 = (3x^3 - x^2 - 6x)\left(\tfrac{1}{3}x + \tfrac{4}{9}\right) + \left(\tfrac{13}{9}x^2 + \tfrac{2}{3}x - 2\right).$$

Therefore $\frac{13}{9}x^2 + \frac{2}{3}x - 2 \in I$.

$$3x^3 - x^2 - 6x = \left(\tfrac{13}{9}x^2 + \tfrac{2}{3}x - 2\right)\left(\tfrac{27}{13}x - \tfrac{9\cdot 31}{13^2}\right) + \left(\tfrac{-2\cdot 3^2\cdot 7}{13^2}x - \tfrac{2\cdot 3^2\cdot 31}{13^2}\right).$$

Therefore $\frac{-2\cdot 3^2\cdot 7}{13^2}x - \frac{2\cdot 3^2\cdot 31}{13^2} \in I$.

$$\tfrac{13}{9}x^2 + \tfrac{2}{3}x - 2 = \left(\tfrac{-2\cdot 3^2\cdot 7}{13^2}x - \tfrac{2\cdot 3^2\cdot 31}{13^2}\right)\left(\tfrac{-13^3}{2\cdot 3^4\cdot 7}x + \tfrac{13^2 19^2}{2\cdot 3^4 7^2}\right) + \left(\tfrac{13^2 61}{3^2 7^2}\right).$$

Therefore $\frac{13^2 61}{3^2 7^2} \in I$. And $\frac{3^2 7^2}{13^2 61} \cdot \frac{13^2 61}{3^2 7^2} = 1$ therefore $1 \in I$. Therefore $I = \mathbb{C}[x]$. Therefore the g.c.d. of $f(x)$ and $g(x)$ equals 1: $\gcd(f, g) = 1$.

Solution 2 (replacing the first polynomial with $2x^5 - x^3 - 3x^2 - 6x + 6$).

Now we'll solve it assuming $f(x) = 2x^5 - x^3 - 3x^2 - 6x + 6$ which is what I think they really intended. Again let $I = \langle f(x), g(x) \rangle$ be the ideal in $\mathbb{C}[x]$ generated by f and g. Dividing g into f we get

$$2x^5 - x^3 - 3x^2 - 6x + 4 = (x^4 + x^3 - x^2 - 2x - 2)(2x - 2) + (3x^3 - x^2 - 6x + 2). \tag{30}$$

Thus $(3x^3 - x^2 - 6x + 2) \in I$. Dividing again

$$x^4 + x^3 - x^2 - 2x - 2 = (3x^3 - x^2 - 6x + 2)\left(\tfrac{1}{3}x + \tfrac{4}{9}\right) + \tfrac{13}{9}(x^2 - 2). \tag{31}$$

Thus $x^2 - 2 \in I$. Dividing again we have

$$3x^3 - x^2 - 6x + 2 = (x^2 - 2)(3x - 1), \tag{32}$$

and the remainder is zero. Now we solve back, substituting (32) into (31) gives

$$x^4 + x^3 - x^2 - 2x - 2 = (x^2 - 2)(3x - 1)\left(\tfrac{1}{3}x + \tfrac{4}{9}\right) + \tfrac{13}{9}(x^2 - 2)$$

and factoring out $x^2 - 2$ from both terms on the right hand side we get

$$x^4 + x^3 - x^2 - 2x - 2 = (x^2 - 2)(x^2 + x + 1). \tag{33}$$

Thus $g \in (x^2 - 2)\mathbb{C}[x]$. Now substituting (32) and (33) into (30) gives

$$2x^5 - x^3 - 3x^2 - 6x + 4 = (x^2 - 2)(x^2 + x + 1)(2x - 2) + (x^2 - 2)(3x - 1)$$

and factoring out $x^2 - 2$ from both terms on the right hand side we get $f \in (x^2 - 2)\mathbb{C}[x]$. Thus we've shown both f and g are in $(x^2 - 2)\mathbb{C}$. Since I equals $\langle f, g\rangle$, it follows that $I \subseteq (x^2 - 2)\mathbb{C}[x]$. We've shown $(x^2 - 2)\mathbb{C}[x] \subseteq I$. Thus we can conclude $(x^2 - 2)\mathbb{C}[x] = I$. It follows that the g.c.d. of f and g is $x^2 - 2$.

(b) Let $f(x) = 3x^4 + 8x^2 - 3$ and $g(x) = x^3 + 2x^2 + 3x + 6$. Let $I = \langle f(x), g(x)\rangle$. We have

$$3x^4 + 8x^2 - 3 = (x^3 + 2x^2 + 3x + 6)(3x - 6) + (11x^2 + 33)$$

Thus $11x^2 + 33 \in I$. Thus $x^2 + 3 \in I$. Now

$$x^3 + 2x^2 + 3x + 6 = (x^2 + 3)(x + 2).$$

Thus $x^2 + 3$ divides both $f(x)$ and $g(x)$. In particular

$$f(x) = (3x^2 - 1)(x^2 + 3)$$

$$g(x) = (x + 2)(x^2 + 3)$$

and therefore it follows that

$$\langle f(x), g(x)\rangle = \langle x^2 + 3\rangle.$$

(c) As in part (a) I have a feeling there's a typo and they really meant $x^3 + 6x^2 + 7x + 2$ because then both are divisible by $x + 1$. But as it is, $x^3 + 6x^2 + 7x + 1$ is irreducible and the calculations are even worse than part (a). As I did in part (a), I'll solve this in both the way they actually stated it and the was I think they intended it. First the way they stated it.

<u>*Solution 1 (as stated in book).*</u>

Let $f(x) = x^4 - 2x^3 - 2x^2 - 2x - 3$ and $g(x) = x^3 + 6x^2 + 7x + 1$ and $I = \langle f(x), g(x)\rangle$.

$$x^4 - 2x^3 - 2x^2 - 2x - 3 = (x^3 + 6x^2 + 7x + 1)(x - 8) + (33x^2 + 53x + 5)$$

Thus $33x^2 + 53x + 5 \in I$.

$$x^3 + 6x^2 + 7x + 1 = (33x^2 + 53x + 5)\left(\tfrac{1}{33}x + \tfrac{145}{1089\cdot 33^2}\right) + \left(\tfrac{-227}{1089}x + \tfrac{364}{1089}\right).$$

Thus

$$\tfrac{-227}{1089}x + \tfrac{364}{1089} \in I$$

And finally

$$33x^2 + 53x + 5 = \left(\tfrac{-227}{1089}x + \tfrac{364}{1089}\right)\left(\tfrac{-1089\cdot 33}{227}x - \tfrac{26182827}{51529}\right) + \left(\tfrac{9009297}{51529}\right).$$

Thus $\frac{9009297}{51529} \in I$ and it follows that $1 \in I$ and that $\gcd(f(x), g(x)) = 1$.

<u>*Solution 2 (replacing the second polynomial with $x^3 + 6x^2 + 7x + 2$).*</u>

Exercise 3: Let A be an $n\times n$ matrix over a field F. Show that the set of all polynomials f in $F[x]$ such that $f(A) = 0$ is an ideal.

Solution: Let I be the set $\{f \in F[x] \mid f(A) = 0\}$. Let $f, g \in I$ and $c \in F$. Then $(cf + g)(A) = cf(A) + g(A) = c\cdot 0 + 0 = 0$. Thus I is a subspace of $F[x]$ as a vector space over F. Now suppose $f \in I$ and $g \in F[x]$. Then $(gf)(A) = g(A)f(A) = g(A)\cdot 0 = 0$. Thus I has the required property to be an ideal.

Exercise 4: Let F be a subfield of the complex numbers, and let

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix}.$$

Find the monic generator of the ideal of all polynomials f in $F[x]$ such that $f(A) = 0$.

Solution: Let $I = \{f(x) \in F[x] \mid f(A) = 0\}$. We know from the previous exercise that I is an ideal. It's clear that $f(A) \neq 0$ if $\deg(f) \leq 1$. Thus if we find any $f(x) \in I$ such that $\deg(f) = 2$ then f must be a generator. Let $f(x) = x^2 - 4x + 3$. Then

$$f(A) = A^2 - 4A + 3$$
$$= \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} - 4 \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & -8 \\ 0 & 9 \end{bmatrix} - \begin{bmatrix} 4 & -8 \\ 0 & 12 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

You will learn an easy way to find this polynomial in Theorem 4 of Section 6.3.

Exercise 5: Let F be a field. Show that the intersection of any number of ideals in $F[x]$ is an ideal.

Solution: Let A be a set. Let I_α be an ideal in $F[x]$ for each $\alpha \in A$. Let $I = \cap_{\alpha \in A} I_\alpha$. By Theorem 2, page 36 (sec. 2.2), I is a subspace of $F[x]$. We must just show I satisfies the extra condition required to be an ideal. Specifically, let $g(x) \in F[x]$ and $f(x) \in I$. We must show $g(x)f(x) \in I$. Since $f(x) \in I$, it follows that $f(x) \in I_\alpha\ \forall\ \alpha \in A$. Since I_α is an ideal $\forall \alpha \in A$, it follows that $g(x)f(x) \in I_\alpha\ \forall\ \alpha \in A$. Thus $f(x)g(x) \in I_\alpha\ \forall\ \alpha \in A$. Thus $f(x)g(x) \in I$.

Exercise 6: Let F be a field. Show that the ideal generated by a finite number of polynomials $f_1, \ldots, f_n$ in $F[x]$ is the intersection of all ideals containing $f_1, \ldots, f_n$.

Solution: Let $I = \langle f_1, \ldots, f_n \rangle$ be the ideal generated by $\{f_1, \ldots, f_n\}$. Let J be any ideal containing $\{f_1, \ldots, f_n\}$. Let $g \in I$. Then $g = g_1 f_1 + \cdots + g_n f_n$ for some $g_1, \ldots, g_n$ in $F[x]$. Since J is an ideal containing $\{f_1, \ldots, f_n\}$ it must also contain g. Thus $g \in I \Rightarrow g \in J$. Thus $I \subseteq J$. Now by the definition of intersection, J is contained in every ideal which contains $\{f_1, \ldots, f_n\}$. Since I is such an ideal, it follows that $J \subseteq I$. Thus we've shown $I \subseteq J$ and $J \subseteq I$. Thus $J = I$.

Exercise 7: Let K be a subfield of a field F, and suppose f, g are polynomials in $K[x]$. Let M_K be the ideal generated by f and g in $K[x]$ and M_F be the ideal they generate in $F[x]$. Show that M_K and M_F have the same monic generator.

Solution:

Solution 1 We first note a general fact. Let E be any field. Let $f, g \in E[x]$ such that $\gcd(f, g) = k$. Let $h \in E[x]$ be monic. We claim $\gcd(kf, kg) = kh$ in $E[x]$. Since $\gcd(f, g) = k$, the ideal generated by f and g is the same as the ideal generated by k. In other words $fE[x] + gE[x] = kE[x]$. Thus $hfE[x] + hgE[x] = hkE[x]$. Since h is monic by assumption and k is monic by definition of gcd, it follows that hk is monic - and therefore hk satisifes the definition of $\gcd(f, g)$ in $E[x]$.

Now let $h = \gcd(f, g)$ in $K[x]$. So $h \in K[x]$ and $h = fa + gb$ for some $a, b \in K[x]$. And $\exists\ u, v \in K[x]$ such that $f = uh$ and $g = vh$. Once you can write the three equations $h = fa + gb$, $f = uh$ and $g = vh$ it follows that the ideal in $K[x]$ generated by f and g is the same as the ideal generated by h. That's what it means for h to be the g.c.d. of f and g. But these three equalities also hold in $F[x]$ and consequently the ideal in $F[x]$ generated by f and g is the same as the one generated by h.

Solution 2 This might be easier if they had formalized what is discussed informally on page 132 after the proof of Theorem 7. What they are describing is the so-called "eucidean algorithm". If you have two polynomials f and g and say $\deg(f) \geq \deg(g)$. Then you divide g into f and take the remainder $r \in K[x]$. That remainder is also in the ideal generated f and g. Then replace f and g with g and r and do the same thing, divide r into g and the remainder has degree less than r. Eventually you must arrive at a remainder equal to zero or of degree equal to zero (i.e. a scalar). If the remainder is zero then the previous remainder generates the ideal $\langle f, g \rangle$ in $K[x]$. If the remainder is a non-zero scalar then the ideal contains 1 and therefore $\langle f, g \rangle = K[x]$.

The euclidean algorithm allows us to always find the gcd in a finite number of steps.

Once we have the euclidean algorithm, all we need to do to solve this problem is to note that all operations in the euclidean algorithm happen in $K[x]$ and so if $K \subseteq F$ then in $F[x]$ the same operations will lead to the same generator which will therefore still live in the smaller ring $K[x]$.

Section 4.5: The Prime Factorization of a Polynomial

Page 137: In the proof of Theorem 11, they definitely use the product and chain rules for derivatives. I don't believe those have ever been proven or even stated. The product rule was actually proven as part of Exercise 9 part (e), from Section 4.2, page 123.

Page 139: In Exercise 7 it says "Use Exercise 7". They meant to say "Use Exercise 6".

Exercise 1: Let p be a monic polynomial over the field F, and let f and g be relatively prime polynomials ovef F. Prove that the g.c.d. of pf and pg is p.

Solution: I proved this in more generality as part of Exercise 7 of the previous section. That makes me think that was not the solution they were looking for. Or maybe they want a proof here using prime factorization. Here are both proofs.

Solution 1 We use the definition of g.c.d. in terms of ideals. Since $\gcd(f, g) = 1$, the ideal generated by f and g is the same as the ideal generated by 1. In other words $fF[x] + gF[x] = F[x]$. Thus $pfF[x] + pgF[x] = pF[x]$. Since p is monic by assumption, it follows that p satisifes the definition of $\gcd(f, g)$ in $F[x]$.

Solution 2 We use the comments at the top of page 137. Factor into primes $f = f_1^{r_1} \cdots f_j^{r_j}$, $g = g_1^{s_1} \cdots g_k^{s_k}$ and $p = p_1^{t_1} \cdots p_\ell^{t_\ell}$. Since $\text{g.c.d.}(f, g) = 1$ none of the f_i's equal any of the g_i's. Thus the common factors of pf and pg are exactly the p_i's. In other words the prime factorizations are

$$fp = f_1^{r_1} \cdots f_j^{r_j} \cdot p_1^{t_1} \cdots p_\ell^{t_\ell}$$

$$gp = g_1^{s_1} \cdots g_k^{s_k} \cdot p_1^{t_1} \cdots p_\ell^{t_\ell}.$$

Since none of the f_i's equal any of the g_i's, it follows that $\text{g.c.d.}(f, g) = p_1^{t_1} \cdots p_\ell^{t_\ell}$ which equals p.

Exercise 2: Assuming the Fundamental Theorem of Algebra prove the following. If f and g are polynomials over the field of complex numbers, then $\text{g.c.d.}(f, g) = 1$ if and only if f and g have no common root.

Solution: By the Fundamental Theorem of Algebra, we can factor f and g all the way to linear factors

$$f = (x - a_1)^{n_1} \cdots (x - a_k)^{n_k}$$

$$g = (x - b_1)^{m_1} \cdots (x - b_\ell)^{m_\ell}.$$

The roots of f are exactly $a_1, \dots, a_k$, the roots of g are exactly $b_1, \dots, b_\ell$. If $\gcd(f, g) = 1$ then (by the comments at the top of page 137) f and g have no common factors, and therefore $a_i \neq b_j \ \forall\ i, j$. Thus (by Corollary 1, page 128) f and g have no common roots. And if f and g have no common roots then (by Corollary 1, page 128) none of the factors $(x - a_i)$ can equal any of the factors $(x - b_j)$. Thus $\gcd(f, g) = 1$.

Exercise 3: Let D be the differentiation operator on the space of polynomials over the field of complex numbers. Let f be a monic polynomial over the field of complex numbers. Prove that

$$f = (x - c_1) \cdots (x - c_k)$$

where $c_1, \ldots, c_k$ are *distinct* complex numbers if and only if f and Df are relatively prime. In other words, f has no repeated roots if and only if f and Df have no common root. (Assume the Fundamental Theorem of Algebra.)

Solution: First assume all c_i's are distinct. Then by Theorem 11, page 137, we know f and Df are relatively prime.

Now assume f and f' are relatively prime. Then again by Theorem 11 we know f is a product of *distinct* irreducibles. Since $\mathbb{C}$ is algebraically closed each of those irreducibles must be of the form $x - a$. It follows that f must be of the form $(x - c_1)\cdots(x - c_n)$ for distinct $c_1, \ldots, c_n$.

Exercise 4: Prove the following generalization of Taylor's formula. Let f, g, and h be polynomials over a subfield of the complex numbers, with $\deg f \le n$. Then

$$f(g) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(h)(g - h)^k.$$

(Here $f(g)$ denotes 'f of g.')

Solution: Since we are working over the base field $\mathbb{C}$ by Theorem 3 page 126 there is a natural isomorphism between the algebra $\mathbb{C}[x]$ and the algebra of polynomial functions from $\mathbb{C}$ into $\mathbb{C}$. Taylor's Theorem is therefore also a theorem about polynomial functions from $\mathbb{C}$ to $\mathbb{C}$. So we may plug any complex numbers we want in for x and c in the statement of Taylor's Theorem and the equality remains valid. In particular we can subsitute g in for x and h in for c to obtain the desired formula. We then simply translate over to the algebra $\mathbb{C}[x]$ via the isomorphism in Theorem 3 to obtain the corresponding result in $\mathbb{C}[x]$.

For the remaining exercises, we shall need the following definition. If f, g, and p are polynomials over the field F with $p \neq 0$, we say f is **congruent to** g **modulo** p if $(f - g)$ is divisible by p. If f is congruent to g modulo p, we write

$$f \equiv g \pmod{p}.$$

Exercise 5: Prove for any non-zero polynomial p, that congruence modulo p is an equivalence relation.

(a) It is reflexive: $f \equiv f \pmod{p}$.

(a) It is symmetric: if $f \equiv g \pmod{p}$, then $g \equiv f \pmod{p}$.

(a) It is transitive: if $f \equiv g \pmod{p}$ and $g \equiv h \pmod{p}$, then $f \equiv h \pmod{p}$.

Solution:

(a). In this case p must divide $f - f$ which equals zero. But everything divides zero. Just take $d = 0$ and then $p = d(f - f)$. Thus $f \equiv f \pmod{p}$ is always true.

(b). Assume $f \equiv g \pmod{p}$. Then p divides $f - g$. Thus $\exists\, d$ s.t. $f - g = pd$. Let $d' = -d$. Then $g - f = pd'$ thus p divides $g - f$. Thus $g \equiv f \pmod{p}$.

(c). Assume $f \equiv g \pmod{p}$ and $g \equiv h \pmod{p}$. Then $\exists\, d$ and d' such that $f - g = pd$ and $g - h = pd'$. Adding these two equations gives $f - h = pd + pd'$. Let $d'' = d + d'$. Then $f - h = pd''$. Thus $f \equiv h \pmod{p}$.

Exercise 6: Suppose $f \equiv g \pmod{p}$ and $f_1 \equiv g_1 \pmod{p}$.

(a) Prove that $f + f_1 \equiv g + g_1 \pmod{p}$.

(b) Prove that $ff_1 \equiv gg_1 \pmod{p}$.

Solution:

(a) By assumption there are polynomials d and d' such that $f - g = pd$ and $f_1 - g_1 = pd'$. Adding these two equations gives $f + f_1 - g - g_1 = pd + pd'$. Or equivalently $(f + f_1) - (g + g_1) = p(d + d')$. Thus $f + f_1 \equiv g + g_1 \pmod{p}$.

(b) By assumption there are polynomials d and d' such that $f - g = pd$ and $f_1 - g_1 = pd'$. Now $ff_1 - gg_1 = ff_1 - g_1 f + g_1 f - gg_1 = f(f_1 - g_1) + g_1(f - g) = fpd' + g_1 pd = p(fd' + g_1 d)$. Thus p divides $ff_1 - gg_1$. Thus $ff_1 \equiv gg_1 \pmod{p}$.

Exercise 7: Use Exercise 6 to prove the following. If f, g, and p are polynomials over the field F and $p \neq 0$, and if $f \equiv g \pmod{p}$, then $h(f) \equiv h(g) \pmod{p}$.

Solution: It follows from Exercise 6 part (b) that $f \equiv g \pmod{p} \Rightarrow f^2 \equiv g^2 \pmod{p}$ (since $f \equiv f \pmod{p}$ and $g \equiv g \pmod{p}$). By induction $f^n \equiv g^n \pmod{p}$ for all $n = 1, 2, 3, \ldots$. Let $c \in F$. Then letting $f_1 = g_1 = c$ we get $cf \equiv cg \pmod{p}$ for any $c \in F$. Thus if $h(x) = cx^n$ we have shown the result is true. Thus the result is true for all monomials. Now we can obtain the result on the sum of monomials using part (a) of Exercise 6. Since the general h is a sum of monomials, the general result follows.

Exercise 8: If p is an irreducible polynomial and $fg \equiv 0 \pmod{p}$, prove that either $f \equiv 0 \pmod{p}$ or $g \equiv 0 \pmod{p}$. Give an example which shows that this is false if p is not irreducible.

Solution: $fg \equiv 0 \pmod{p}$ implies p divides fg. By the Corollary to Theorem 8, page 135, it follows that p divides f of p divides g. Thus $f \equiv 0 \pmod{p}$ or $g \equiv 0 \pmod{p}$.

Chapter 5: Determinants

Section 5.2: Determinant Functions

Exercise 1: Each of the following expressions defines a function D on the set of 3×3 matricces over the field of real numbers. In which of these cases is D a 3-linear function?

(a) $D(A) = A_{11} + A_{22} + A_{33}$;

(b) $D(A) = (A_{11})^2 + 3A_{11}A_{22}$;

(c) $D(A) = A_{11}A_{12}A_{33}$;

(d) $D(A) = A_{13}A_{22}A_{32} + 5A_{12}A_{22}A_{32}$;

(e) $D(A) = 0$;

(f) $D(A) = 1$;

Solution:

(a) No D is not 3-linear. Let

$$A = \left[\begin{array}{ccc} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

Then if D were 3-linear then it would be linear in the first row and we'd have to have $D(A) = D(I) + D(I)$. But $D(A) = 4$ and $D(I) = 3$, so $D(A) \neq D(I) + D(I)$.

(b) No D is not 3-linear. Let A be the same matrix as in part (a). Then $D(A) = 10$ and $D(I) = 4$, so $D(A) \neq D(I) + D(I)$.

(c) No D is not 3-linear. Let

$$A = \left[\begin{array}{ccc} 2 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right],$$

$$B = \left[\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

Then if D were 3-linear we'd have to have $D(A) = D(B) + D(B)$. But $D(A) = 4$ and $D(B) = 1$. Thus $D(A) \neq D(B) + D(B)$.

(d) Yes D is 3-linear. The two functions $A \mapsto A_{13}A_{22}A_{32}$ and $A \mapsto 5A_{12}A_{22}A_{32}$ are both 3-linear by Example 1, page 142. The sum of these two functions is then 3-linear by the Lemma on page 143. Since D is exactly the sum of these two functions, it follows that D is 3-linear.

(e) Yes D is 3-linear. We must show (5-1) on page 142 holds for all matrices A. But since $D(A) = 0\ \forall\ A$, both sides of (5-1) are always equal to zero. Thus (5-1) does hold $\forall\ A$.

(f) No D is not 3-linear. Let A be the matrix from part (a) again. Then $D(A) = 1$ but $D(I)+D(I) = 2$. Thus $D(A) \neq D(I)+D(I)$. Thus D is not 3-linear.

Exercise 2: Verify directly that the three functions E_1, E_2, E_3 defined by (5-6), (5-7), and (5-8) are identical.

Solution:

$$\begin{aligned}E_1(A) &= A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{21}(A_{12}A_{33} - A_{13}A_{32}) + A_{31}(A_{12}A_{23} - A_{13}A_{22})\\ &= \underset{\text{term 1}}{A_{11}A_{22}A_{33}} - \underset{\text{term 2}}{A_{11}A_{23}A_{32}} - \underset{\text{term 3}}{A_{21}A_{12}A_{33}} + \underset{\text{term 4}}{A_{21}A_{13}A_{32}} + \underset{\text{term 5}}{A_{31}A_{12}A_{23}} - \underset{\text{term 6}}{A_{31}A_{13}A_{22}}.\end{aligned}$$

$$\begin{aligned}E_2(A) &= -A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{22}(A_{11}A_{33} - A_{13}A_{31}) - A_{32}(A_{11}A_{23} - A_{13}A_{21})\\ &= -\underset{\text{term 3}}{A_{12}A_{21}A_{33}} + \underset{\text{term 5}}{A_{12}A_{23}A_{31}} + \underset{\text{term 1}}{A_{22}A_{11}A_{33}} - \underset{\text{term 6}}{A_{22}A_{13}A_{31}} - \underset{\text{term 2}}{A_{32}A_{11}A_{23}} + \underset{\text{term 4}}{A_{32}A_{13}A_{21}}.\end{aligned}$$

$$\begin{aligned}E_3(A) &= A_{13}(A_{21}A_{32} - A_{22}A_{31}) - A_{23}(A_{11}A_{32} - A_{12}A_{31}) + A_{33}(A_{11}A_{22} - A_{12}A_{21})\\ &= \underset{\text{term 4}}{A_{13}A_{21}A_{32}} - \underset{\text{term 6}}{A_{13}A_{22}A_{31}} - \underset{\text{term 2}}{A_{23}A_{11}A_{32}} + \underset{\text{term 5}}{A_{23}A_{12}A_{31}} + \underset{\text{term 1}}{A_{33}A_{11}A_{22}} - \underset{\text{term 3}}{A_{33}A_{12}A_{21}}.\end{aligned}$$

I've expanded the three expressions and labelled corresponding terms. We see each of the six terms appears exactly once in each expansion, and always with the same sign. Therefore the three expressions are equal.

Exercise 3: Let K be a commutative ring with identity. If A is a 2×2 matrix over K, the **classical adjoint** of A is the 2×2 matrix adj A defined by

$$\text{adj } A = \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}.$$

If det denotes the unique determinant function on 2×2 matrices over K, show that

(a) $(\text{adj } A)A = A(\text{adj } A) = (\det A)I$;

(a) $\det(\text{adj } A) = \det(A)$;

(a) $\text{adj } (A^t) = (\text{adj } A)^t$.

(A^t denotes the transpose of A.)

Solution:

(a)

$$\begin{aligned}(\text{adj } A)A &= \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\\ &= \begin{bmatrix} A_{11}A_{22} - A_{12}A_{21} & A_{12}A_{22} - A_{12}A_{22} \\ -A_{11}A_{21} + A_{11}A_{21} & -A_{12}A_{21} + A_{11}A_{22} \end{bmatrix}\\ &= \begin{bmatrix} A_{11}A_{22} - A_{12}A_{21} & 0 \\ 0 & A_{11}A_{22} - A_{12}A_{21} \end{bmatrix}\end{aligned}$$

$$= \begin{bmatrix} \det(A) & 0 \\ 0 & \det(A) \end{bmatrix}.$$

$$\begin{aligned} A(\text{adj } A) &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}A_{22} - A_{12}A_{21} & -A_{11}A_{12} + A_{12}A_{11} \\ A_{21}A_{22} - A_{22}A_{21} & -A_{21}A_{12} + A_{22}A_{11} \end{bmatrix} \\ &= \begin{bmatrix} \det(A) & 0 \\ 0 & \det(A) \end{bmatrix}. \end{aligned}$$

Thus both (adj A)A and A(adj A) equal $(\det A)I$.

(b)

$$\begin{aligned} \det(\text{adj } A) &= \det\left(\begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}\right) \\ &= A_{11}A_{22} - A_{12}A_{21} \\ &= \det(A). \end{aligned}$$

(c)

$$\begin{aligned} \text{adj}(A^t) &= \text{adj}\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^t\right) \\ &= \text{adj}\left(\begin{bmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{bmatrix}\right) \\ &= \begin{bmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{bmatrix} \end{aligned} \tag{34}$$

And

$$\begin{aligned} (\text{adj}A)^t &= \left(\text{adj}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\right)^t \\ &= \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}^t \\ &= \begin{bmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{bmatrix} \end{aligned} \tag{35}$$

Comparing (34) and (35) gives the result.

Exercise 4: Let A be a 2×2 matrix over a field F. Show that A is invertible if and only if $\det A \neq 0$. When A is invertible, give a formula for A^{-1}.

Solution: We showed in Example 3, page 143, that $\det(A) = A_{11}A_{22} - A_{12}A_{21}$. Therefore, we've already done the first part in Exercise 8 of section 1.6 (page 27). We just need a formula for A^{-1}. The formula is

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}.$$

Checking:

$$A \cdot \frac{1}{\det(A)} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}.$$

$$= \frac{1}{\det(A)} \begin{bmatrix} A_{11}A_{22} - A_{12}A_{21} & -A_{11}A_{12} + A_{12}A_{11} \\ A_{21}A_{22} - A_{22}A_{21} & -A_{21}A_{12} + A_{22}A_{11} \end{bmatrix}$$

$$= \frac{1}{\det(A)} \begin{bmatrix} \det(A) & 0 \\ 0 & \det(A) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Exercise 5: Let A be a 2×2 matrix over a field F, and suppose that $A^2 = 0$. Show for each scalar c that $\det(cI - A) = c^2$.

Solution: One has to be careful in proving this not to use implications such as $2x = 0 \Rightarrow x = 0$; or $x^2 + y = 0 \Rightarrow y = 0$. These implications are not valid in a general field. However, we will need to use that fact that $xy = 0 \Rightarrow x = 0$ or $y = 0$, which is true in any field.

Let $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$. Then

$$A^2 = \begin{bmatrix} x^2 + yz & xy + yw \\ xz + wz & yz + w^2 \end{bmatrix}.$$

If $A^2 = 0$ then

$$x^2 + yz = 0 \tag{36}$$

$$y(x + w) = 0 \tag{37}$$

$$z(x + w) = 0 \tag{38}$$

$$yz + w^2 = 0. \tag{39}$$

Now $\det(cI - A) = \det \begin{bmatrix} c - x & -y \\ -z & c - w \end{bmatrix} = (c - x)(c - w) - yz = c^2 - c(x + w) + xw - yz.$

Thus

$$\det(cI - A) = c^2 - c(x + w) + \det(A). \tag{40}$$

Suppose $x + w \neq 0$. Then (37) and (38) imply $y = z = 0$. Thus $A = \begin{bmatrix} x & 0 \\ 0 & w \end{bmatrix}$. But then $A^2 = \begin{bmatrix} x^2 & 0 \\ 0 & w^2 \end{bmatrix}$. So if $A^2 = 0$ then it must be that also $x = w = 0$, which contradicts the assumption that $x + w \neq 0$.

Thus necessarily $A^2 = 0$ implies $x + w = 0$. This implies $A = \begin{bmatrix} x & y \\ z & -x \end{bmatrix}$. Thus $\det(A) = -x^2 - yz$, which equals zero by (36). Thus $A^2 = 0$ implies $x + w = 0$ and $\det(A) = 0$. Thus by (40) $A^2 = 0$ implies $\det(cI - A) = c^2$.

Exercise 6: Let K be a subfield of the complex numbers and n a positive integer. Let $j_1, \ldots, j_n$ and $k_1, \ldots, k_n$ be positive integers not exceeding n. For an $n \times n$ matrix A over K define

$$D(A) = A(j_1, k_1)A(j_2, k_2) \cdots A(j_n, k_n).$$

Prove that D is n-linear if and only if the integers $j_1, \ldots, j_n$ are distinct.

Solution: First assume the integers $j_1, \ldots, j_n$ are distinct. Since these n integers all satisfy $1 \leq j_i \leq n$, their being distinct necessarily implies $\{j_1, \ldots, j_n\} = \{1, 2, 3, \ldots, n\}$ Thus $A(j_1, k_1)A(j_2, k_2) \cdots A(j_n, k_n)$ is just a rearrangement of $A(1, k_1)A(2, k_2) \cdots A(3, k_n)$. It follows from Example 1 on page 142 that $A(j_1, k_1)A(j_2, k_2) \cdots A(j_n, k_n)$ is n-linear.

Now assume two or more of the j_i's are equal. Assume without loss of generality that $j_1 = j_2 = \cdots = j_\ell = 1$ where $\ell \geq 2$. Let A be the matrix with all 2's in the first row and all ones in all other rows. Let B be the matrix of all 1's. Then $D(A) = 2^\ell$ and $D(B) = 1$. Since D is n-linear it must be that $D(A) = D(B) + D(B)$. But $\ell > 1 \Rightarrow 2^\ell \neq 2$. Thus $D(A) \neq D(B) + D(B)$ and

D is not n-linear.

Exercise 7: Let K be a commutative ring with identity. Show that the determinant function on 2×2 matrices A over K is alternating and 2-linear as a function of the columns of A.

Solution:

$$\det\begin{bmatrix} ra_1+a_2 & b \\ rc_1+c_2 & d\end{bmatrix}$$
$$=(ra_1+a_2)d-(rc_1+c_2)b$$
$$=ra_1d+a_2d-rc_1b-c_2b$$
$$=r(ad-bc_1)+(a_2d-bc_2)$$
$$=r\det\begin{bmatrix} a_1 & b \\ c_1 & d\end{bmatrix}+\det\begin{bmatrix} a_2 & b \\ c_2 & d\end{bmatrix}.$$

Likewise

$$\det\begin{bmatrix} a & rb_1+b_2 \\ c & rd_1+d_2\end{bmatrix}$$
$$=a(rd_1+d_2)-(rb_1+b_2)c$$
$$=rad_1+ad_2-rcb_1-cb_2$$
$$=(rad_1-rcb_1)+(ad_2-cb_2)$$
$$=r\det\begin{bmatrix} a & b_1 \\ c & d_1\end{bmatrix}+\det\begin{bmatrix} a & b_2 \\ c & d_2\end{bmatrix}.$$

Thus the determinant function is 2-linear on columns. Now

$$\det\begin{bmatrix} a & b \\ c & d\end{bmatrix}$$
$$=ad-bc$$
$$=-(bc-ad)$$
$$=-\det\begin{bmatrix} b & d \\ a & c\end{bmatrix}.$$

Thus the determinant function is alternating on columns.

Exercise 8: Let K be a commutative ring with identity. Define a function D on 3×3 matrices over K by the rule

$$D(A)=A_{11}\det\begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33}\end{bmatrix}-A_{12}\begin{bmatrix} A_{21} & A_{23} \\ A_{31} & A_{33}\end{bmatrix}+A_{13}\begin{bmatrix} A_{21} & A_{22} \\ A_{31} & A_{32}\end{bmatrix}.$$

Show that D is alternating and 3-linear as a function of the columns of A.

Solution: This is exactly Theorem 1 page 146 but with respect to columns instead of rows. The statement and proof go through without change except for chaning the word "row" to "column" everywhere. To make it work, however, we must know that det is an alternating 2-linear function on *columns* of 2×2 matrices over K. This is exactly what was shown in the previous exercise.

Exercise 9: Let K be a commutative ring with identity and D and alternating n-linear function on $n\times n$ matrices over K. Show that

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

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

Solution: Let A be an $n \times n$ matrix with one row all zeros. Suppose row α_i is all zeros. Then $\alpha_i + \alpha_i = \alpha_i$. Thus by the linearity of the determinant in the i^{th} row we have $\det(A) = \det(A) + \det(A)$. Subtracting $\det(A)$ from both sides gives $\det(A) = 0$.

Now suppose B is obtained from A by adding a scalar multiple of one row to another. Assume row β_i of B equals $\alpha_i + c\alpha_j$ where α_i is the i^{th} row of A and α_j is the j^{th}. Then the rows of B are $\alpha_1, \ldots, \alpha_{i-1}, \alpha_i + c\alpha_j, \alpha_{i+1}, \ldots, \alpha_n$. Thus

$$\det(B) = \det(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i + c\alpha_j, \alpha_{i+1}, \ldots, \alpha_n)$$

$$= \det(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i + c\alpha_j, \alpha_{i+1}, \ldots, \alpha_n)$$

$$= \det(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_n) + c \cdot \det(\alpha_1, \ldots, \alpha_{i-1}, \alpha_j, \alpha_{i+1}, \ldots, \alpha_n).$$

The first determinant is $\det(A)$. And by the first part of this problem, the second determinant equals zero, since it has a repeated row α_j. Thus $\det(B) = \det(A)$.

Exercise 10: Let F be a field, A a 2×3 matrix over F, and (c_1, c_2, c_3) the vector in F^3 defined by

$$c_1 = \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix}, \quad c_2 = \begin{vmatrix} A_{13} & A_{11} \\ A_{23} & A_{21} \end{vmatrix}, \quad c_3 = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}.$$

Show that

(a) $\operatorname{rank}(A) = 2$ if and only if $(c_1, c_2, c_3) \neq 0$;

(b) if A has rank 2, then (c_1, c_2, c_3) is a basis for the solution space of the system of equations $AX = 0$.

Solution: We will use the fact that the rank of a 2×2 matrix is 2 $\Leftrightarrow$ the matrix is invertible $\Leftrightarrow$ the determinant is non-zero. The first equivalence follows from the fact that a matrix M with rank 2 has two linearly independent rows and therefore the row space of M is all of F^2 which is the same as the row space of the identity matrix. Thus by the Corollary on page 58 M is row-equivalent to the identity matrix, thus by Theorem 12 (page 23) it follows that M is invertible. The second equivalence follows from Exercise 4 from Section 5.2 (page 149).

(a) If $\operatorname{rank}(A) = 0$ then A is the zero matrix and clearly $c_1 = c_2 = c_3 = 0$.

If $\operatorname{rank}(A) = 1$ then the second row must be a multiple of the first row. This is then true for each of the 2×2 matrices

$$\begin{bmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{bmatrix}, \quad \begin{bmatrix} A_{13} & A_{11} \\ A_{23} & A_{21} \end{bmatrix}, \quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{41}$$

because each one is obtained from A by deleting one column (and in the case of the second one, switching the two remaining columns). Thus each of them has rank ≤ 1. Therefore the determinant of each of these three matrices is zero. Thus (c_1, c_2, c_3) is the zero vector.

If $\operatorname{rank}(A) = 2$ then the second row of A is not a multiple of its first row. We must show the same is true of at least one of the matrices in (41). Suppose the second row is a multiple of the first for each matrix in (41). Since each pair of these matrices shares a column, it must be the same multiple for each pair; and therefore the same multiple for all three, call it c. Therefore the seond row of the entire matrix A is c times the first row, which contradicts our assumption that $\operatorname{rank}(A) = 2$. Thus at least one of the matrices in (41) must have rank two and the result follow.

(b) Identify F^3 with the space of 3×1 column vectors and F^2 the space of 2×1 column vectors. Let $T : F^3 \to F^2$ be the linear transformation given by $TX = AX$. Then by Theorem 2 page 71 (the rank/nullity theorem) we know $\operatorname{rank}(T) + \operatorname{nullity}(T) = 3$. It was shown in the proof of Theorem 3 page 72 (the third displayed equation in the proof) that $\operatorname{rank}(T) = \text{column rank}(A)$. And $\operatorname{nullity}(T)$ is the solution space for $AX = 0$. Thus column $\operatorname{rank}(A)$ plus the rank of the solution space of $AX = 0$ equals three. Thus if rank(A) = 2 then the rank of the solution space of $AX = 0$ must equal one. Thus a basis for this space is any non-zero vector in the space. Thus we only need show (c_1, c_2, c_3) is in this space. In other words we must show

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0.$$

It feels like we're supposed to apply Exercise 8 to the following matrix

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix},$$

but the problem with that is that we do not know that an alternating function on columns is necessarily zero on a matrix with a repeated *row*. That is true, but rather than prove it, it's easier just prove this directly

$$c_1 = A_{12}A_{23} - A_{22}A_{13}$$

$$c_1 = A_{13}A_{21} - A_{11}A_{23}$$

$$c_1 = A_{11}A_{22} - A_{12}A_{21}.$$

Therefore

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}c_1 + A_{12}c_2 + A_{13}c_3 \\ A_{21}c_1 + A_{22}c_2 + A_{23}c_3 \end{bmatrix}$$

Expanding the first entry

$$A_{11}c_1 + A_{12}c_2 + A_{13}c_3$$

$$= A_{11}(A_{12}A_{23} - A_{22}A_{13}) + A_{12}(A_{13}A_{21} - A_{11}A_{23}) + A_{13}(A_{11}A_{22} - A_{12}A_{21})$$

$$= A_{11}A_{12}A_{23} - A_{11}A_{22}A_{13} - A_{12}A_{13}A_{21} - A_{12}A_{11}A_{23} + A_{13}A_{11}A_{22} - A_{13}A_{12}A_{21}$$

matching up terms we see everything cancels.

$$= \underbrace{A_{11}A_{12}A_{23}}_{\text{term 1}} - \underbrace{A_{11}A_{22}A_{13}}_{\text{term 2}} - \underbrace{A_{12}A_{13}A_{21}}_{\text{term 3}} - \underbrace{A_{12}A_{11}A_{23}}_{\text{term 1}} + \underbrace{A_{13}A_{11}A_{22}}_{\text{term 2}} - \underbrace{A_{13}A_{12}A_{21}}_{\text{term 3}} = 0.$$

Expanding the second entry

$$A_{21}c_1 + A_{22}c_2 + A_{23}c_3$$

$$= A_{21}(A_{12}A_{23} - A_{22}A_{13}) + A_{22}(A_{13}A_{21} - A_{11}A_{23}) + A_{23}(A_{11}A_{22} - A_{12}A_{21})$$

$$= A_{21}A_{12}A_{23} - A_{21}A_{22}A_{13} + A_{22}A_{13}A_{21} - A_{22}A_{11}A_{23} + A_{23}A_{11}A_{22} - A_{23}A_{12}A_{21}$$

matching up terms we see everything cancels.

$$= \underbrace{A_{21}A_{12}A_{23}}_{\text{term 1}} - \underbrace{A_{21}A_{22}A_{13}}_{\text{term 2}} + \underbrace{A_{22}A_{13}A_{21}}_{\text{term 2}} - \underbrace{A_{22}A_{11}A_{23}}_{\text{term 3}} + \underbrace{A_{23}A_{11}A_{22}}_{\text{term 3}} - \underbrace{A_{23}A_{12}A_{21}}_{\text{term 1}} = 0.$$

Section 5.3: Permutations and the Uniqueness of Determinants

Exercise 9: Let n be a positive integer and F a field. If σ is a permutation of degree n, prove that the function

$$T(x_1, \ldots, x_n) = (x_{\sigma 1}, \ldots, x_{\sigma n})$$

is an invertible linear operator on F^n.

Solution:

Chapter 6: Elementary Canonical Forms

Section 6.2: Characteristic Values

Exercise 1: In each of the following cases, let T be the linear operator on $\mathbb{R}^2$ which is represented by the matrix A in the standard ordered basis for $\mathbb{R}^2$, and let U be the linear operator on $\mathbb{C}^2$ represented by A in the standard ordered basis. Find the characyteristic polynomial for T and that for U, find the characteristic values of each operator, and for each such charactersistic value c find a basis for the corresponding space of characteristic vectors.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Solution:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$xI - A = \begin{bmatrix} x-1 & 0 \\ 0 & x \end{bmatrix}$$

The characteristic polynomial equals $|xI - A| = x(x-1)$. So $c_1 = 0$, $c_1 = 1$. A basis for W_1 is $\{(0,1)\}$, $\alpha_1 = (0,1)$. A basis for W_2 is $\{(1,0)\}$, $\alpha_2 = (1,0)$. This is the same whether the base field is $\mathbb{R}$ or $\mathbb{C}$ since the characteristic polynomial factors completely.

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

$$|xI - A| = \begin{vmatrix} x-2 & -3 \\ 1 & x-1 \end{vmatrix}$$

$= (x-2)(x-1) + 3 = x^2 - 3x + 5$. This is a parabola opening up with vertex $(3/2, 11/4)$. Thus there are no real roots. Using the quadratic formula $c_1 = \frac{3+\sqrt{11}i}{2}$ and $c_2 = \frac{3-\sqrt{11}i}{2}$. To find the a characteristic vector for c_1 we solve

$$\begin{bmatrix} \frac{-1+\sqrt{11}i}{2} & 3 \\ -1 & \frac{1+\sqrt{11}i}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives the characteristic vector $\alpha_1 = (\frac{1+\sqrt{11}i}{2}, 1)$. To find the a characteristic vector for c_2 we solve

$$\begin{bmatrix} \frac{-1-\sqrt{11}i}{2} & 3 \\ -1 & \frac{1-\sqrt{11}i}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives the characteristic vector $\alpha_2 = (\frac{1-\sqrt{11}i}{2}, 1)$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$|xI - A| = \begin{vmatrix} x-1 & 1 \\ 1 & x-1 \end{vmatrix} = (x-1)^2 - 1 = x(x-2)$. So $c_1 = 0$ for which $\alpha_1 = (1, -1)$. And $c_2 = 2$ for which $\alpha_2 = (1, 1)$. This is the same in both $\mathbb{R}$ and $\mathbb{C}$ since the characteristic polynomial factors completely.

Exercise 2: Let F be an n-dimensional vector space over F. What is the characteristic polynomial of the identity operator on V? What is the characteristic polynomial for the zero operator?

Solution: The identity operator can be represented by the $n \times n$ identity matrix I. The characteristic polynomial of the identity operator is therefore $(x-1)^n$. The zero operator is represented by the zero matrix in any basis. Thus the characteristic polynomial of the zero operator is x^n.

Exercise 3: Let A be an $n \times n$ triangular matrix over the field F. Prove that the characteristic values of A are the diagonal entries of A, i.e., the scalars A_{ii}.

Solution: The determinant of a triangular matrix is the product of the diagonal entries. Thus $|xI - A| = \prod(x - a_{ii})$.

Exercise 4: Let T be the linear operator of $\mathbb{R}^3$ which is represented in the standard ordered basis by the matrix

$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}.$$

Prove that T is diagonalizable by exhibiting a basis for $\mathbb{R}^3$, each vector fo which is a characteristic vector of T.

Solution:

$$|xI - A| = \begin{vmatrix} x+9 & -4 & -4 \\ 8 & x-3 & -4 \\ 16 & -8 & x-7 \end{vmatrix}$$

$$= \begin{vmatrix} x+9 & 0 & -4 \\ 8 & x+1 & -4 \\ 16 & -x-1 & x-7 \end{vmatrix}$$

$$= (x+1)\begin{vmatrix} x+9 & 0 & -4 \\ 8 & 1 & -4 \\ 16 & -1 & x-7 \end{vmatrix}$$

$$= (x+1)\begin{vmatrix} x+9 & 0 & -4 \\ 8 & 1 & -4 \\ 24 & 0 & x-11 \end{vmatrix}$$

$$= (x+1)\begin{vmatrix} x+9 & -4 \\ 24 & x-11 \end{vmatrix}$$

$= (x+1)[(x+9)(x-11)+96] = (x+1)(x^2-2x-3) = (x+1)(x-3)(x+1) = (x+1)^2(x-3)$. Thus $c_1 = -1$, $c_2 = 3$. For c_1, $xI - A$ equals

$$= \begin{bmatrix} 8 & -4 & -4 \\ 8 & -4 & -4 \\ 16 & -8 & -8 \end{bmatrix}$$

This matrix evidently has rank one. Thus the null space has rank two. The two characteristic vectors $(1,2,0)$ and $(1,0,2)$ are independent, so they form a basis for W_1. For c_2, $xI-A$ equals

$$= \begin{bmatrix} -12 & -4 & -4 \\ 8 & 0 & -4 \\ 16 & -8 & -4 \end{bmatrix}$$

This is row equivalent to

$$= \begin{bmatrix} 1 & -0 & -1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus the null space one dimensional and is given by $(z/2, z/2, z)$. So $(1,1,2)$ is a characteristic vector and a basis for W_2. By Theorem 2 (ii) T is diagonalizable.

Exercise 5: Let

$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

Is A similar over the field $\mathbb{R}$ to a diagonal matrix? Is A similar over the field $\mathbb{C}$ to a diagonal matrix?

Solution:

$$= \begin{vmatrix} x-6 & 3 & 2 \\ -4 & x+1 & 2 \\ -10 & 5 & x+3 \end{vmatrix}$$

$$= \begin{vmatrix} x-6 & 3 & 2 \\ -x+2 & x-2 & 0 \\ -10 & 5 & x+3 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-6 & 3 & 2 \\ -1 & 1 & 0 \\ -10 & 5 & x+3 \end{vmatrix}$$

$$= (x-2)\begin{vmatrix} x-3 & 3 & 2 \\ 0 & 1 & 0 \\ -5 & 5 & x+3 \end{vmatrix}$$

$= (x-2)((x-3)(x+3)+10) = (x-2)(x^2+1)$. Since this is not a product of linear factors over $\mathbb{R}$, by Theorem 2, page 187, A is not diagonalizable over $\mathbb{R}$. Over $\mathbb{C}$ this factors to $(x-2)(x-i)(x+i)$. Thus over $\mathbb{C}$ the matrix A has three distinct characteristic values. The space of characteristic vectors for a given characteristic value has dimension at least one. Thus the sum of the dimensions of the W_i's must be at least n. It cannot be greater than n so it must equal n exactly. Thus A is diagonalizable over $\mathbb{C}$.

Exercise 6: Let T be the linear operator on $\mathbb{R}^4$ which is represented in the standard ordered basis by the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{bmatrix}.$$

Under what conditions on a, b, and c is T diagonalizable?

Solution:

$$|xI-A| = \begin{vmatrix} x & 0 & 0 & 0 \\ -a & x & 0 & 0 \\ 0 & -b & x & 0 \\ 0 & 0 & -c & x \end{vmatrix}$$

$= x^4$. Therefore there is only one characteristic value $c_1 = 0$. Thus $c_1 I - A = A$ and W_1 is the null space of A. So A is diagonalizable $\Leftrightarrow \dim(W) = 4 \Leftrightarrow A$ is the zero matrix $\Leftrightarrow a = b = c = 0$.

Exercise 7: Let T be the linear operator on the n-dimensional vector space V, and suppose that T has n *distinct* characteristic values. Prove that T is diagonalizable.

Solution: The space of characteristic vectors for a given characteristic value has dimension at least one. Thus the sum of the dimensions of the W_i's must be at least n. It cannot be greater than n so it must equal n exactly. Thus by Theorem 2, T is diagonalizable.

Exercise 8: Let A and B be $n \times n$ matrices over the field F. Prove that if $(I - AB)$ is invertible, then $I - BA$ is invertible and

$$(I - BA)^{-1} = I + B(I - AB)^{-1}A.$$

Solution:

$$\begin{aligned}
&(I - BA)(I + B(I - AB)^{-1}A) \\
&= I - BA + B(I - AB)^{-1}A - BAB(I - AB)^{-1}A \\
&= I - B(A - (I - AB)^{-1}A + AB(I - AB)^{-1}A) \\
&= I - B(I - (I - AB)^{-1} + AB(I - AB)^{-1}A \\
&= I - B(I - (I - AB)(I - AB)^{-1})A \\
&= I - B(I - I)A \\
&= I.
\end{aligned}$$

Exercise 9: Use the result of Exercise 8 to prove that, if A and B are $n \times n$ matrices over the field F, then AB and BA have precisely the same characteristic values in F.

Solution: By Theorem 3, page 154, $\det(AB) = \det(A)\det(B)$. Thus AB is singular $\Leftrightarrow$ BA is singular. Therefore 0 is a characteristic values of $AB \Leftrightarrow 0$ is a characteristic value of BA. Now suppose the characteristic value c of AB is not equal to zero. Then $|cI - AB| = 0 \Leftrightarrow c^n|I - \frac{1}{c}AB| = 0 \overset{\text{by \#8}}{\Leftrightarrow} c^n|I - \frac{1}{c}BA| = 0 \Leftrightarrow |cI - BA| = 0$.

Exercise 10: Suppose that A is a 2×2 matrix with real entries which is symmetrix ($A^t = A$). Prove that A is similar over $\mathbb{R}$ to a diagonal matrix.

Solution: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. So $|xI - A| = \begin{vmatrix} x-a & -b \\ -b & x-a \end{vmatrix} = (x-a)^2 - b^2 = (x-a-b)(x-a+b)$. So $c_1 = a+b$, $c_2 = a = b$. If $b = 0$ then A is already diagonal. If $b \neq 0$ then $c_1 \neq c_2$ so by Exercise 7 A is diagonalizable.

Exercise 11: Let N be a 2×2 complex matrix such that $N^2 = 0$. Prove that either $N = 0$ or N is similar over $\mathbb{C}$ to

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

Solution: Suppose $N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Now $N^2 = 0 \Rightarrow \begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}$ are characteristic vectors for the characteristic value 0. If $\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}$ are linearly independent then W_1 has rank two and N is diagonalizable to $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. If $PNP^{-1} = 0$ then $N = P^{-1}0P = 0$ so in this case N itself is the zero matrix. This contradicts the assumption that $\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}$ are linearly independent.

So we can assume that $\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}$ are linearly dependent. If both equal the zero vector then $N = 0$. So we can assume at least one vector is non-zero. If $\begin{bmatrix} b \\ d \end{bmatrix}$ is the zero vector then $N = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}$. So $N^2 = 0 \Rightarrow a^2 = 0 \to a = 0$. Thus $N = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}$. In this case N is similar to $N = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ via the matrix $P = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$. Similary if $\begin{bmatrix} a \\ c \end{bmatrix}$ is the zero vector, then $N^2 = 0$ implies $d^2 = 0$ implies $d = 0$ so $N = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$. In this case N is similar to $N = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}$ via the matrix $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is simiilar to $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ as above.

By the above we can assume neither $\begin{bmatrix} a \\ c \end{bmatrix}$ or $\begin{bmatrix} b \\ d \end{bmatrix}$ is the zero vector. Since they are linearly dependent we can assume $\begin{bmatrix} b \\ d \end{bmatrix} = x\begin{bmatrix} a \\ c \end{bmatrix}$ so $N = \begin{bmatrix} a & ax \\ c & cx \end{bmatrix}$. So $N^2 = 0$ implies

$$a(a + cx) = 0$$

$$c(a + cx) = 0$$

$$ax(a + cx) = 0$$

$$cx(a + cx) = 0.$$

We know that at least one of a or c is not zero. If $a = 0$ then since $c \neq 0$ it must be that $x = 0$. So in this case $N = \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}$ which is similar to $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ as before. If $a \neq 0$ then $x \neq 0$ else $a(a + cx) = 0$ implies $a = 0$. Thus $a + cx = 0$ so $N = \begin{bmatrix} a & ax \\ -a/x & -a \end{bmatrix}$. This is similar to $\begin{bmatrix} a & a \\ -a & -a \end{bmatrix}$ via $P = \begin{bmatrix} \sqrt{x} & 0 \\ 0 & 1/\sqrt{x} \end{bmatrix}$. And $\begin{bmatrix} a & a \\ -a & -a \end{bmatrix}$ is similar to $\begin{bmatrix} 0 & 0 \\ -a & 0 \end{bmatrix}$ via $P = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$. And this finally is similar to $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ as before.

Exercise 12: Use the result of Exercise 11 to prove the following: If A is a 2×2 matrix with complex entries, then A is similar over $\mathbb{C}$ to a matrix of one of the two types

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \qquad \begin{bmatrix} a & 0 \\ 1 & a \end{bmatrix}.$$

Solution: Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Since the base field is $\mathbb{C}$ the characteristic polynomial $p(x) = (x - c_1)(x - c_2)$. If $c_1 \neq c_2$ then A is diagonalizable by Exercise 7. If $c_1 = c_2$ then $p(x) = (x - c_1)^2$. If W has dimension two then A is diagonalizable by Theorem 2. Thus we will be done if we show that if $p(x) = (x - c_1)^2$ and $\dim(W_1) = 1$ then A is similar to $\begin{bmatrix} a & 0 \\ 1 & a \end{bmatrix}$.

We will need the following three identities:

$$\begin{bmatrix} a & 0 \\ c & d \end{bmatrix} \sim \begin{bmatrix} a & 0 \\ 1 & d \end{bmatrix} \quad \text{via } p = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \tag{42}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \sim \begin{bmatrix} a-b & c-d \\ b & d \end{bmatrix} \quad \text{via } p = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \tag{43}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \sim \begin{bmatrix} a & xb \\ c/x & d \end{bmatrix} \quad \text{via } p = \begin{bmatrix} \sqrt{x} & 0 \\ 0 & 1/\sqrt{x} \end{bmatrix} \text{for } x \neq 0. \tag{44}$$

Now we know in this case that A is not diagonalizable. If $d \neq 0$ then $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \sim \begin{bmatrix} a & bc/d \\ d & d \end{bmatrix}$ by (44) with $x = c/d$ and this in turn is similar to $\begin{bmatrix} a - bc/d & 0 \\ -a + 2bc/d & d \end{bmatrix}$ by (43).

Now we know the diagonal entries are the characteristic values, which are equal. Thus $a - \frac{bc}{d} = d$. So this equals $\begin{bmatrix} d & 0 \\ x & d \end{bmatrix}$ where $x = -a + \frac{2bc}{d}$ and we know $x \neq 0$ since A is not diagonalizable. Thus $A \sim \begin{bmatrix} d & 0 \\ 1 & d \end{bmatrix}$ by (42). Now suppose $d = 0$. Then $A = \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & c \\ b & a \end{bmatrix}$ via $p = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. If $b = 0$ then $A = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}$ and again since A has equal characteristic values it must be that $a = 0$. So $A = \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}$ which is similar to $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ via $P = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$. So assume $b \neq 0$. Then $A \sim \begin{bmatrix} 0 & c \\ b & a \end{bmatrix}$ and we can argue exact as above were $d \neq 0$.

Exercise 13: Let V be the vector space of all functions from $\mathbb{R}$ into $\mathbb{R}$ which are continuous, i.e., the space of continuous real-valued functions on the real line. Let T be the linear operator on V defined by

$$(Tf)(x) = \int_0^x f(t)dt.$$

Prove that T has no characteristic values.

Solution: Suppose $\exists\, c$ such that $Tf = cf\ \forall\, f$. Then $\int_0^x f(t)dt = cf(x)$. Let $f(x) = 1$. Then we must have $\int_0^x 1dt = c \cdot 1$, which imples $x = c$. In other words this implies the functions $f(x) = x$ and $g(x) = c$ are the same, which they are not because one is constant and the other is not. This therefore is a contradiction. Thus it is impossible that $\exists\, c$ such that $Tf = cf\ \forall\, f$. Thus T has no characteristic values.

Exercise 14: Let A be an $n \times n$ *diagonal* matrix with characteristic polynomial

$$(x - c_1)^{d_1} \cdots (x - c_k)^{d_k},$$

where $c_1, \ldots, c_k$ are distinct. Let V be the space of $n \times n$ matrices B such that $AB = BA$. Prove that the dimension of V is $d_1^2 + \cdots + d_k^2$.

Solution: Write

$$A = \begin{bmatrix} c_1 I & & & \\ & c_2 I & 0 & \\ & 0 & \ddots & \\ & & & c_k I \end{bmatrix}.$$

Write

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1k} \\ B_{21} & B_{22} & \cdots & B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ B_{k1} & B_{k2} & \cdots & B_{kk} \end{bmatrix}$$

where B_{ij} has dimenson $d_i \times d_j$. Then $AB = BA$ implies

$$\begin{bmatrix} c_1 B_{11} & c_1 B_{12} & \cdots & c_1 B_{1k} \\ c_2 B_{21} & c_2 B_{22} & \cdots & c_2 B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_k B_{kk} & c_k B_{k2} & \cdots & c_k B_{kk} \end{bmatrix} = \begin{bmatrix} c_1 B_{11} & c_2 B_{12} & \cdots & c_k B_{1k} \\ c_1 B_{21} & c_2 B_{22} & \cdots & c_k B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 B_{k1} & c_2 B_{k2} & \cdots & c_k B_{kk} \end{bmatrix}$$

Thus $c_i \neq c_j$ for $i \neq j$ implies $B_{ij} = 0$ for $i \neq j$, while $B_{11}, B_{22}, \ldots, B_{kk}$ can be arbitrary. The dimension of B_{ii} is therefore d_i^2 thus the dimension of the space of all such B_{ii}'s is $d_1^2 + d_2^2 + \cdots + d_k^2$.

Exercise 15: Let V be the space of $n \times n$ matrices over F. Let A be a fixed $n \times n$ matrix over F. Let T be the linear operator 'left multiplication by A' on V. Is it true that A and T have the same characteristic values?

Solution: Yes. Represent an element of V as a column vector by stacking the columns of V on top of each other, with the first column on top. Then the matrix for T is given by $\left[\begin{array}{cccc} A & & & \\ & A & 0 & \\ & 0 & \ddots & \\ & & & A \end{array}\right]$. By the argument on page 157 the determinant of this matrix is $\det(A)^n$. Thus if p is the characteristic polynomial of A then p^n is the characteristic polynomial of T. Thus they have exactly the same roots and thus they have exactly the same characteristic values.

Section 6.3: Annihilating Polynomials

Page 198: Typo in Exercise 11, "Section 6.1" should be "Section 6.2".

Exercise 1: Let V be a finite-dimensional vector space. What is the minimal polynomial for the identity operator on V? What is the minimal polynomial for the zero operator?

Solution: The minimal polynomial for the identity operator is $x - 1$. It annihilates the identity operator and the monic zero degree polynomial $p(x) = 1$ does not, so it must be the minimal polynomial. The minimal polynomial for the zero operator is x. It is a monic polynomial that annihilates the zero operator and again the monic zero degree polynomial $p(x) = 1$ does not, so it must be the minimal polynomial.

Exercise 2: Let a, b and c be tlements of a field F, and let A be the following 3×3 matrix over F:

$$A = \left[\begin{array}{ccc} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{array}\right].$$

Prove that the characteristic polynomial for A is $x^x - ax^2 - bx - c$ and that this is also the minimal polynomial for A.

Solution: The characteristic polynomial is

$$\left|\begin{array}{ccc} x & 0 & -c \\ -1 & x & -b \\ 0 & -1 & x-a \end{array}\right| = \left|\begin{array}{ccc} x & 0 & -c \\ -1 & 0 & x^2 - ax - b \\ 0 & -1 & x - a \end{array}\right| = 1 \cdot \left|\begin{array}{cc} x & -c \\ -1 & x^2 - ax - b \end{array}\right| = x^3 - ax^2 - bx - c.$$

Now for any $r, s \in F$

$$A^2 + rA + s = \left[\begin{array}{ccc} 0 & c & ac \\ 0 & b & c + ba \\ 1 & a & b + a^2 \end{array}\right] + \left[\begin{array}{ccc} 0 & 0 & rc \\ r & 0 & rb \\ 0 & r & ra \end{array}\right] + \left[\begin{array}{ccc} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{array}\right]$$

$$= \left[\begin{array}{ccc} s & c & ac + rc \\ r & b + s & c + ba + br \\ 1 & a + r & b + a^2 + ra + s \end{array}\right] \neq 0.$$

Thus $f(A) \neq 0$ for all $f \in F[x]$ such that $\deg(F) = 2$. Thus the minimum polynomial cannot have degree two, it must therefore have degree three. Since it divides $x^3 - ax^2 - bx - c$ it must equal $x^3 - ax^2 - bx - c$.

Exercise 3: Let A be the 4×4 real matrix

$$\left[\begin{array}{cccc} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{array}\right].$$

Show that the characteristic polynomial for A is $x^2(x-1)^2$ and that it is also the minimal polynomial.

Solution: The characteristic polynomial equals

$$\left|\begin{array}{cccc} x-1 & -1 & 0 & 0 \\ 1 & x+1 & 0 & 0 \\ 2 & 2 & x-2 & -1 \\ -1 & -1 & 1 & x \end{array}\right| = \left|\begin{array}{cc} x-1 & -1 \\ 1 & x+1 \end{array}\right| \cdot \left|\begin{array}{cc} x-2 & -1 \\ 1 & x \end{array}\right| \quad \text{by (5-20) page 158}$$

$$= x^2(x^2-2x+1) = x^2(x-1)^2.$$

The minimum polynomial is clearly not linear, thus the minimal polynomial is one of $x^2(x-1)^2$, $x^2(x-1)$, $x(x-1)^2$ or $x(x-1)$. We will plug A in to the first three and show it is not zero. It will follow that the minimum polynomial must be $x^2(x-1)^2$.

$$A^2 = \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -3 & -3 & 3 & 2 \\ 2 & 2 & -2 & -1 \end{array}\right]$$

$$A - I = \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 \\ -2 & -2 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{array}\right]$$

and

$$(A-I)^2 = \left[\begin{array}{cccc} -1 & -2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus

$$A^2(A-I) = \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{array}\right] \neq 0$$

$$A(A-I)^2 = \left[\begin{array}{cccc} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \neq 0$$

and

$$A(A-I) = \left[\begin{array}{cccc} -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -2 & -2 \end{array}\right] \neq 0.$$

Thus the minimal polynomial must be $x^2(x-1)^2$.

Exercise 4: Is the matrix A of Exercise 3 similar over the field of complex numbers to a diagonal matrix?

Solution: Not diagonalizable, because for characteristic value $c = 0$ the matrix $A - cI = A$ and A is row equivalent to

$$\left[\begin{array}{cccc}1&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{array}\right]$$

which has rank three. So the null space has dimension one. So if W is the null space for $A - cI$ then W has dimension one, which is less than the power of x in the characteristic polynomial. So by Theorem 2, page 187, A is not diagonalizable.

Exercise 5: Let V be an n-dimensional vector space and let T be a linear operator on V. Suppose that there exists some positive integer k so that $T^k = 0$. Prove tht $T^n = 0$.

Solution: $T^k = 0 \Rightarrow$ the only characteristic value is zero. We know the minimal polynomial divides this so the minimal polynomial is of the form t^r for some $1 \leq r \leq n$. Thus by Theorem 3, page 193, the characteristic polynomial's only root is zero, and the characteristic polynomial has degree n. So the characteristic polynomial equals t^n. By Theorem 4 (Caley-Hamilton) $T^n = 0$.

Exercise 6: Find a 3×3 matrix for which the minimal polynomial is x^2.

Solution: If $A^2 = 0$ and $A \neq 0$ then the minimal polynomial is x or x^2. So any $A \neq 0$ such that $A^2 = 0$ has minimal polynomial x^2. E.g.

$$A = \left[\begin{array}{ccc}0&0&0\\1&0&0\\0&0&0\end{array}\right].$$

Exercise 7: Let n be a positive integer, and let V be the space of polynomials over $\mathbb{R}$ which have degree at most n (throw in the 0-polynomial). Let D be the differentiation operator on V. What is the minimal polynomial for D?

Solution: $1, x, x^2, \ldots, x^n$ is a basis.

$$1 \mapsto 0$$

$$x \mapsto 1$$

$$x^2 \mapsto 2x$$

$$\vdots$$

$$x^n \mapsto nx^{n-1}$$

The matrix for D is therefore

$$\left[\begin{array}{cccccc}0&1&0&0&\cdots&0\\0&0&2&0&\cdots&0\\0&0&0&3&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&0&\cdots&n\end{array}\right]$$

Suppose A is a matrix such that $a_{ij} = 0$ except when $j = i + 1$. Then A^2 has $a_{ij} = 0$ except when $j = i + 2$. A^3 has $a_{ij} = 0$ except when $j = i + 3$. Etc., where finally $A^n = 0$. Thus if $a_{ij} \neq 0 \;\forall\; j = i + 1$ then $A^k \neq 0$ for $k < n$ and $A^n = 0$. Thus the minimum polynomial divides x^n and cannot be x^k for $k < n$. Thus the minimum polynomial is x^n.

Exercise 8: Let P be the operator on $\mathbb{R}^2$ which projects each vector onto the x-axis, parallel to the y-axis: $P(x, y) = (x, 0)$. Show that P is linear. What is the minimal polynomial for P?

Solution: P can be given in the standard basis by left multiplication by $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Since P is given by left multiplication by a matrix, P is clearly linear. Since A is diagonal, the characteristic values are the diagonal values. Thus the characteristic values of A are 0 and 1. The characteristic polynomial is a degree two monic polynomial for which both 0 and 1 are roots. Therefore the characteristic polynomial is $x(x-1)$. If the characteristic polynomial is a product of distinct linear terms then it must equal the minimal polynomial. Thus the minimal polynomial is also $x(x-1)$.

Exercise 9: Let A be an $n \times n$ matrix with characteristic polynomial

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}.$$

Show that

$$c_1 d_1 + \cdots + c_k d_k = \text{trace}(A).$$

Solution: Suppose A is $n \times n$. Claim: $|xI - A| = x^n + \text{trace}(A)x^{n-1} + \cdots$. Proof by induction: case $n = 2$. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. $|xI - A| = x^2 + (a+d)x + (ad - bc)$. The trace of A is $a + d$ so we have established the claim for the case $n = 2$. Suppose true for up to $n - 1$. Let $r = a_{22} + a_{33} + \cdots + a_{nn}$. Then

$$\begin{vmatrix} x - a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & x - a_{22} & \cdots & a_{2n} \\ \ddots & \ddots & \cdots & \ddots \\ a_{n1} & a_{n2} & \cdots & x - a_{nn} \end{vmatrix}$$

Now expanding by minors using the first column, and using induction, we get that this equals

$$(x - a_{11})(x^{n-1} - rx^{n-2} + \cdots)$$

$$-a_{21}(\text{polynomial of degree } n-2)$$

$$+a_{31}(\text{polynomial of degree } n-2)$$

$$+\cdots$$

$$= x^n + (r + a_{11})x^{n-1} + \text{polynomial of degree at most } n-2$$

$$= x^n - \text{tr}(A)x^{n-1} + \cdots$$

Now if $f(x) = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$ then the coefficient of x^{n-1} is $c_1 d_1 + \cdots c_k d_k$ so it must be that $c_1 d_1 + \cdots c_k d_k = \text{tr}(A)$.

Exercise 10: Let V be the vector space of $n \times n$ matrices over the field F. Let A be a fixed $n \times n$ matrix. Let T be the linear operator on V defined by

$$T(B) = AB.$$

Show that the minimal polynomial for T is the minimal polynnomial for A.

Solution: If we represent a $n \times n$ matrix as a column vector by stacking the columns of the matrix on top of each other, with the first column on the top, then the transformation T is represented in the standard basis by the matrix

$$M = \begin{bmatrix} A & & & \\ & A & 0 & \\ & 0 & \ddots & \\ & & & A \end{bmatrix}.$$

And since

$$f(M) = \begin{bmatrix} f(A) & & & \\ & f(A) & 0 & \\ & 0 & \ddots & \\ & & & f(A) \end{bmatrix}$$

it is evident that $f(M) = 0 \Leftrightarrow f(A) = 0$.

Exercise 11: Let A and B be $n \times n$ matrices over the field F. According to Exercise 9 of Section 6.2, the matrices AB and BA have the same characteristic values. Do they have the same characteristic polynomial? Do they have the same minimal polynomial?

Solution: In Exercise 9 Section 6.2 we showed $|xI = AB| = 0 \Leftrightarrow |xI - BA| = 0$. Thus we have two monic polynomials of degree n with exactly the same roots. Thuse they are equal. So the characteristic polynomials are equal. But the minimum polynomials need not be equal. To see this let $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $AB = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ so the minimal polynomial of BA is x and the minimal polynomial of AB is clearly not x (it is in fact x^2).

Section 6.4: Invariant Subspaces

Exercise 1: Let T be the linear operator on $\mathbb{R}^2$, the matrix of which in the standard ordered basis is

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

(a) Prove that the only subspaces of $\mathbb{R}^2$ invariant under T are $\mathbb{R}^2$ and the zero subspace.

(b) If U is the linear operator on $\mathbb{C}^2$, the matrix of which in the standard ordered basis is A, show that U has 1-dimensional invariant subspaces.

Solution: (a) The charactersistic polynomial equals $\begin{vmatrix} x-1 & 1 \\ -2 & x-2 \end{vmatrix} = (x-1)(x-2)+2 = x^2-3x+4$. This is a parabola opening upwards with vertex $(3/2, 7/4)$, so it has no real roots. If T had an invariant subspace it would have to be 1-dimensional and T would therefore have a characteristic value.

(b) Over $\mathbb{C}$ the characteristic polylnomial factors into two linears. Therefore over $\mathbb{C}$, T has two characteristic values and therefore has at least one characteristic vector. The subspace generated by a characteristic vector is a 1-dimensional subspace.

Exercise 2: Let W be an invariant subspace for T. Prove that the minimal polynomial for the restriction operator T_W divides the minimal polynomial for T, without referring to matrices.

Solution: The minimum polynomial of T_W divides any polynomial $f(t)$ where $f(T_W) = 0$. If f is the minimum polynomial for T then $F(T)v = 0\ \forall\ v \in V$. Therefore, $f(T)w = 0\ \forall\ w \in W$. So $f(T_W)w = 0\ \forall\ w \in W$ since by definition $f(T_W)w = f(T)w$ for $w \in W$. Therefore, $f(T_W) = 0$. Therefore the minimum polynomial for T_W divides f.

Exercise 3: Let c be a characteristic value of T and let W be the space of characteristic vectors associated with the characteristic value c. What is the restriction operator T_W?

Solution: For $w \in W$ the transformation $T(w) = cw$. Thus T_W is diagonalizable with single characteristic value c. In other words under which it is represented by the matrix

$$\begin{bmatrix} c & & & \\ & c & 0 & \\ & 0 & \ddots & \\ & & & c \end{bmatrix}$$

where there are $\dim(W)$ c's on the diagonal.

Exercise 4: Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -2 & 2 \\ 2 & -3 & 2 \end{bmatrix}.$$

Is A similar over the field of real numbers to a triangular matrix? If so, find such a triangular matrix.

Solution:

$$A^2 = \begin{bmatrix} 2 & -2 & 2 \\ 0 & 0 & 0 \\ -2 & 2 & -2 \end{bmatrix}.$$

And $A^3 = 0$. Thus the minimal polynomial x^3 and the only characteristic value is 0. We now follow the constructive proof of Theorem 5. $W = \{0\}$, α_1 a characteristic vector of A is $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$. We need α_2 such that $A\alpha_2 \in \{\alpha_1\}$. $\alpha_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ satisfies $A\alpha 2 = \alpha_1$. Now need α_3 such that $A\alpha_3 \in \{\alpha_1, \alpha_2\}$. $\alpha_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ satisfies $A\alpha_3 = 2\alpha_1 + 2\alpha_2$. Thus with respect to the basis $\{\alpha_1, \alpha_2, \alpha_3\}$ the transformation corresponding to A is $\begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

Exercise 5: Every matrix A such that $A^2 = A$ is siimilar to a diagonal matrix.

Solution: $A^2 = A \Rightarrow A$ satisfies the polynomial $x^2 - x = x(x-1)$. Therefore the minimum polynomial of A is either x, $x-1$ or $x(x-1)$. In all three cases the minimum polynomial factors into distinct linears. Therefore, by Theorem 6 A is diagonalizable.

Exercise 6: Let T be a diagonalizable linear opeartor on the n-dimensional vector space V, and let W be a subspace which is invariant under T. Prove that the restriction operator T_W is diagonalizable.

Solution: By the lemma on page 80 the minimum polynomial for T_W divides the minimum polynomial for T. Now T diagonalizable implies (by Theorem 6) that the minimum polynomial for T factors into distinct linears. Since the minimum polynomial for T_W divides it, it must also factor into distinct linears. Thus by Theorem 6 again T_W is diagonalizable.

Exercise 7: Let T be a linear operator on a finite-dimensional vector space over the field of complex numbers. Prove that T is diagonalizable if and only if T is annihilated by some polynomial over $\mathbb{C}$ which has distinct roots.

Solution: If T is diagonalizable then its minimum polynomial is a product of distinct linear factors, and the minimal polynomial annihilates T. This proves "$\Rightarrow$". Now suppose T is annihilated by a polynomial over $\mathbb{C}$ with distinct roots. Since the base field is $\mathbb{C}$ this polynomial factors completely into distinct linear factors. Since the minimum polynomial divides this polynomial the minimum polynomial factors completely into distinct linear factors. Thus by Theorem 6, T is diagonalizable.

Exercise 8: Let T be a linear operator on V. If every subspace of V is invariant under T, then T is a scalar multiple of the identity operator.

Solution: Let $\{\alpha_i\}$ be a basis. The subspace generated by α_i is invariant thus $T\alpha_i$ is a multiple of α_i. Thus α_i is a characteristic vector since $T\alpha_i = c_i\alpha_i$ for some c_i. Suppose $\exists\ i, j$ such that $c_i \neq c_j$. Then $T(\alpha_i + \alpha_j) = T\alpha_i + T\alpha_j = c_i\alpha_i + c_j\alpha_j = c(\alpha_i + \alpha_j)$. Since the subspace generated by $\{\alpha_i, \alpha_j\}$ is invariant under T. Thus $c_i = c$ and $c_j = c$ since coefficients of linear combinations of basis vectors are unique. Thus $T\alpha_i = c\alpha_i\ \forall\ i$. Thus T is c times the identity operator.

Exercise 9: Let T be the indefinite inntegral operator

$$(Tf)(x) = \int_0^x f(t)dt$$

on the space of continuous functions on the interval $[0, 1]$. Is the space of polynomial functions invariant under T? Ths space of differentiable functions? The space of functions which vanish at $x = 1/2$?

Solution: The integral from 0 to x of a polynomial is again a polynomial, so the space of polynomial functions is invariant under T. The integral from 0 to x of a differntiable function is differentiable, so the space of differentiable functions is invariant under T. Now let $f(x) = x - 1/2$. Then f vanishes at $1/2$ but $\int_0^x f(t)dt = \frac{1}{2}x^2 - \frac{1}{2}x$ which does not vanish at $x = 1/2$. So the space of functions which vanish at $x = 1/2$ is not invariant under T.

Exercise 10: Let A be a 3×3 matrix with real entries. Prove that, if A is not similar over $\mathbb{R}$ to a triangular matrix, then A is similar over $\mathbb{C}$ to a diagonal matrix.

Solution: If A is not similar to a tirangular matrix then the minimum polynomial of A must be of the form $(x - c)(x^2 + ax + b)$ where $x^2 + ax + b$ has no real roots. The roots of $x^2 + ax + b$ are then two non-real complex conjugates z and $\bar{z}$. Thus over $\mathbb{C}$ the minimum polynomial factors as $(x - c)(x - z)(x - \bar{z})$. Since c is real, c, z and $\bar{z}$ constintute three distinct numbers. Thus by Theorem 6 A is diagonalizable over $\mathbb{C}$.

Exercise 11: True or false? If the triangular matrix A is similar to a diagonal matrix, then A is already diagonal.

Solution: False. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. Then A is triangular and not diagonal. The characteristic polynomial is $x(x - 1)$ which has distinct roots, so the minimum polynomial is $x(x - 1)$. Thus by Theorem 6, A is diagonalizable.

Exercise 12: Let T be a linear operator on a finite-dimensional vector space over an algebraically closed field F. Let f be a polynomial over F. Prove that c is a characteristic value of $f(T)$ if and only if $c = f(t)$, where t is a characteristic value of T.

Solution: Since F is algebraically closed, the corollary at the bottom of page 203 implies there's a basis under which T is represented by a triangular matrix A. $A = [a_{ij}]$ where $a_{ij} = 0$ if $i > j$ and the a_{ii}, $i = 1, \ldots, n$ are the characteristic values of T. Now $f(A) = [b_{ij}]$ where $b_{ij} = 0$ if $i > j$ and $b_{ii} = f(a_{ii})$ for all $i = 1, \ldots, n$. Thus the characteristic values of $f(A)$ are exactly the $f(c)$'s where c is a characteristic value of A. Since $f(A)$ is a matrix representative of $f(T)$ in the same basis, we conclude the same thing about the tranformation T.

Exercise 13: Let V be the space of $n \times n$ matrices over F. Let A be a fixed $n \times n$ matrix over F. Let T and U be the linear operators on V defined by

$$T(B) = AB$$
$$U(B) = AB - BA$$

(a) True or false? If A is a diagonalizable (over F), then T is diagonalizable.
(b) True or false? If A is diagonalizable, then U is diagonalizable.

Solution: (a) True by Exercise 10 Section 6.3 page 198 since by Theorem 6 diagonalizability depends entirely on the minimum polynomial.

(b) True. Find a basis so that A is diagonal

$$A = \begin{bmatrix} c_1 & & & \\ & c_2 & 0 & \\ & 0 & \ddots & \\ & & & c_n \end{bmatrix}.$$

Let $B = [b_{ij}]$. Then $U(B) = AB - BA$. The n^2 matrices B_{ij} such that $b_{ij} \neq 0$ and all other entries equal zero form a basis for V. For any B_{ij}, $U(B_{ij}) = AB_{ij} - B_{ij}A = [d_{i'j'}]$ where $d_{i'j'} = c_{i'}b_{i'j'} - c_{j'}b_{i'j'} = (c_{i'} - c_{j'})b_{i'j'}$. Thus $d_{i'j'} \neq 0$ only when $i' = i$ and $j' = j$. Thus $U(B_{ij}) = (c_i - c_j)B_{ij}$. So $c_i - c_j$ is a characteristic value and B_{ij} is a characteristic vector for all i, j. Thus V has a basis of characteristic vectors for U. Thus U is diagonalizable.

Section 6.5: Simultaneous Triangulation; Simultaneous Diagonliazation

Exercise 1: Find an invertible real matrix P such that $P^{-1}AP$ and $P^{-1}BP$ are both diagonal, where A and B are the real matrices

(a) $$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -8 \\ 0 & -1 \end{bmatrix}$$

(b) $$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}.$$

Solution: The proof of Theorem 8 shows that if a 2×2 matrix has two characteristic values then the P that diagonalizes A will necessarily also diagonalize any B that commutes with A.

(a) Characteristic polynomial equals $(x-1)(x-2)$. So $c_1 = 1$, $c_2 = 2$.

$$c_1: \quad \begin{bmatrix} 0 & -2 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$c_2: \quad \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $P = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$.

$$P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$P^{-1}BP = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

(b) Characteristic polynomial equals $x(x-2)$. So $c_1 = 0$, $c_2 = 2$.

$$c_1: \quad \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$c_2: \quad \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$.

$$P^{-1}AP = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

$$P^{-1}BP = \begin{bmatrix} 1-a & 0 \\ 0 & 1+a \end{bmatrix}$$

Exercise 2: Let $\mathcal{F}$ be a commuting family of 3×3 complex matrices. How many linearly independent matrices can $\mathcal{F}$ contain? What about the $n \times n$ case?

Solution: This turns out to be quite a hard question, so I'm not sure what Hoffman & Kunze had in mind. But there's a general theorem from 1905 by I. Schur which says the answer is $\left\lfloor \frac{n^2}{4} \right\rfloor + 1$. A simpler proof was published in 1998 by M. Mirzakhani in the American Mathematical Monthly. I have extracted the proof for the case $n = 3$ (which required also extracting the proof for the case $n = 2$).

First we show:

The maximum size of a set of linearly independent commuting triangulizable 2×2 matrices is two $\quad (*)$

Suppose that $\{A_1, A_2, A_3\}$ are three linearly independent commuting upper-triangular 2×2 matrices. Let V be the space generated by $\{A_1, A_2, A_3\}$. So $\dim(V) = 3$.

Write $A_i = \left[\frac{N_i}{0\ M_i}\right]$ where N_i is 1×2. Since $\dim(V) = 3$ it cannot be that all three M_i's are zero. Assume WLOG that $M_1 \neq 0$. Then $M_2 = c_2 M_1$ and $M_3 = c_3 M_1$ for some constants c_1, c_2. Let $B_2 = A_2 - c_2 A_1$ and $B_3 = A_3 = c_3 A_1$. Then $\{B_2, B_3\}$ are lineraly independent in V.

Write $B_2 = \left[\frac{t_2}{0\ 0}\right]$ and $B_3 = \left[\frac{t_3}{0\ 0}\right]$ where $\{t_2, t_3\}$ are linearly independent 1×2 matrices.

Similarly $\exists\ B_2'$ and B_3' in V such that $B_2' = \left[\begin{smallmatrix}0\\0\end{smallmatrix} \middle| t_2'\right]$, $B_3' = \left[\begin{smallmatrix}0\\0\end{smallmatrix} \middle| t_3'\right]$, where $\{t_2', t_3'\}$ are linearly independent.

Since B_2, B_3, B_2', B_3' are all in V, they all commute with each other. Thus $t_i t_j' = 0\ \forall\ i, j$.

Let A be the 2×2 matrix $\left[\begin{smallmatrix}t_3\\t_4\end{smallmatrix}\right]$. Then $\text{rank}(A) = 2$ but $At_2' = 0$ and $At_3' = 0$ thus $\text{null}(A) = 2$. Therefore $\text{rank}(A) + \text{null}(A) = 4$. But $\text{rank}(A) + \text{null}(A)$ cannot be greater than $\dim(V) = 2$. This contradiction imples we cannot have $\{A_1, A_2, A_3\}$ all three be commuting linearly independent upper-triangular 2×2 matrices. But we know we can have two commuting linearly independent upper-triangular 2×2 matrices because $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ are such a pair.

We now turn to the case $n = 3$. Suppose $\mathcal{F}$ is a commuting family of linearly independent 3×3 matrices with $|\mathcal{F}| = 4$. We know $\exists\ P$ such that $P^{-1}\mathcal{F}P$ is a family of upper tirangular commuting matrices. Let V be the space generated by $\mathcal{F}$. Then $\dim(V) = 4$. Let A_1, A_2, A_3, A_4 be a linearly independent subset of V. For each $i\ \exists$ a 2×2 matrix M_i and a 1×3 matrix N_i such that

$$A = \left[\begin{array}{c|c} \multicolumn{2}{c}{N_i} \\ \hline \begin{smallmatrix}0\\0\\0\end{smallmatrix} & M_i \end{array}\right]$$

Since the A_i's commute, for $1 \leq i, j \leq 4$ we have $M_i M_j = M_j M_i$. Suppose W is the vector space spanned by the set $\{M_1, M_2, M_3, M_4\}$ and let $k = \dim(W)$. We know by $(*)$ that $k \leq 2$. Since $\{A_1, A_2, A_3, A_4\}$ are independent we also know $k \geq 1$.

First assume $k = 1$. Then WLOG assume M_1 generates W. Then for $i = 2, 3, 4\ \exists\ n_i$ such that $M_i = n_i M_1$. For $i = 2, 3, 4$ define $B_i = A_i - n_i A_1$. Since $\{A_1, A_2, A_3, A_4\}$ are linearly independent, $\{B_2, B_3, B_4\}$ are linearly independent and $B_i = \left[\begin{smallmatrix}t_i\\0\end{smallmatrix}\right]$ where t_i is 1×3 and $\{t_2, t_3, t_4\}$ are linearly independent.

Now assume $k = 2$. Then WLOG assume M_1, M_2 generate W. Then for each $i = 3, 4\ \exists\ n_{i1}, n_{i2}$ such that $M_i = n_{i1}M_1 + n_{i2}M_2$. For $i = 3, 4$ define $B_i = A_i - n_{i1}A_1 - n_{i2}A_2$. Then $\{A_1, A_2, A_3, A_4\}$ linearly independent implies $\{B_3, B_4\}$ are linearly independent and for $i = 3, 4$, $B_i = \left[\begin{smallmatrix}t_i\\0\end{smallmatrix}\right]$ where t_i is $1 \times n$. Since $\{B_3, B_4\}$ are linearly independent, $\{t_3, t_4\}$ are linearly independent.

Thus in both cases ($k = 1, 2$) we have produced a set of $4 - k$ linearly independent $1 \times n$ matrices $\{t_i\}$ such that $B_i = \left[\begin{smallmatrix} t_i \\ 0 \end{smallmatrix}\right]$.

By a similar argument we obtian a set of two or three linearly independent $n \times 1$ matrices $\{t'_3, t'_4\}$ or $\{t'_2, t'_3, t'_4\}$ such that $B'_i = [0 \mid t'_i]$ is a matrix in V.

Now since all B_i's and B'_i's all belong to the commuting family V, one sees that $t_i t'_j = 0 \; \forall \; i, j$.

Let A be the $m \times 4$ matrix ($m = 2$ or 3) such that its ith row is t_i Since the t_i's are independent we have $\text{rank}(A) \geq m \geq 2$. On the other hand $At'_j = 0$ for all j and the t'_j's are linearly independent. Thus the null space of A has rank greater or equal to the numnber of t'_j's. Thus $\text{rank}(A) \geq 2$ and $\text{null}(A) \geq 2$. But since A is 3×3 we know that $\text{rank}(A) + \text{null}(A) = 3$. This contradiction implies the set $\{A_1, A_2, A_3, A_4\}$ cannot be linearly independent.

Now we can achieve three independent such matrices because $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are such a triple.

Exercise 3: Let T be a linear operator on an n-dimensional space, and suppose that T has n distinct characteristic values. Prove that any linear operator which commutes with T is a polynomial in T.

Solution: Since T has n distinct characteristic values, T is diagonalizable (exercise 6.2.7, page 190). Choose a basis $\mathcal{B}$ for which T is represented by a diagonal matrix A. Suppose the linear transformation S commutes with T. Let B be the matrix of S in the basis $\mathcal{B}$. Then the ij-th entry of AB is $a_{ii}b_{ij}$ and the ij-th entry of BA is $a_{jj}b_{ij}$. Therefore if $a_{ii}b_{ij} = a_{jj}b_{ij}$ and $a_{ii} \neq a_{jj}$, then it must be that $b_{ij} = 0$. So we have shown that B must also be diagonal. So we have to show there exists a polynomial such that $f(a_{ii}) = b_{ii}$ for all $i = 1, \ldots, n$. By Section 4.3 there exists a polynomial with this property.

Exercise 4: Let A, B, C, and D be $n \times n$ complex matrices which commute. Let E be the $2n \times 2n$ matrix

$$E = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Prove that $\det E = \det(AD - BC)$.

Solution: By the corollary on paeg 203 we know A, B, C, and D are all triangulable. By Theorem 7 page 207 we know they are simultaneously triangulable. Let P be the matrix that simultaneously triangulates them. Let

$$M = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix}.$$

Then

$$M^{-1} = \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix}.$$

And

$$M^{-1}EM = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix},$$

where A', B', C', and D' are upper triangular. Now $\det(E) = \det(M^{-1}EM)$. Suppose the result were true for upper triangular matrices A, B, C, and D. Then $\det(E) = \det(M^{-1}EM) = \det(P^{-1}AP \cdot P^{-1}DP - P^{-1}BP \cdot P^{-1}CP) = \det(P^{-1}ADP - P^{-1}BCP) = \det(P^{-1}(AD - BC)P) = \det(AD - BC)$.

Thus it suffices to prove the result for upper triangular matrices. So in what follows we drop the primes and assume A, B, C, and D are upper triangular. We proceed by induction. Suppose first that $n = 1$. Then the theorem is clearly true. Suppose it is true up to $n - 1$.

If A, B, C, and D are upper triangular then it is clear that $\det(AD - BC) = \prod_{i=1}^{n}(a_{ii}d_{ii} - b_{ii}c_{ii})$. So by induction we assume $\det(E) = \prod_{i=1}^{m}(a_{ii}d_{ii} - b_{ii}c_{ii})$ whenever E has dimension $2m$ for $m < n$ (of couse always assuming A, B, C, and D commute).

Now we can write E in the following form for some A', B', C' and D':

$$E = \left[\begin{array}{ccc|ccc} a_{11} & & A' & b_{11} & & B' \\ & \ddots & & & \ddots & \\ 0 & & a_{nn} & 0 & & b_{nn} \\ \hline c_{11} & & C' & d_{11} & & D' \\ & \ddots & & & \ddots & \\ 0 & & c_{nn} & 0 & & d_{nn} \end{array}\right]$$

Expanding E by cofactors of the 1st column gives

$$\det(E) = a_{11}\left|\begin{array}{ccc|cccc} a_{22} & & A'' & 0 & b_{22} & & B'' \\ & \ddots & & & & \ddots & \\ 0 & & a_{nn} & & 0 & & b_{nn} \\ \hline c_{22} & & C' & d_{11} & & D' & \\ & \ddots & & & \ddots & & \\ 0 & & c_{nn} & 0 & & d_{nn} & \end{array}\right| + (-1)^n c_{11}\left|\begin{array}{ccc|cccc} a_{22} & & A' & b_{11} & & B' & \\ & \ddots & & & \ddots & & \\ 0 & & a_{nn} & 0 & & b_{nn} & \\ \hline c_{22} & & C'' & 0 & d_{22} & & D'' \\ & \ddots & & & & \ddots & \\ 0 & & c_{nn} & & 0 & & d_{nn} \end{array}\right|$$

for some A'', B'', C'' and D''.

The $n + 1$ column of each of these matrices has only one non-zero element. So we next expand by cofactors of the $n + 1$-th column of each matrix, which gives

$$(-1)^{2n}a_{11}d_{11}\left|\begin{array}{ccc|ccc} a_{22} & & A'' & b_{22} & & B'' \\ & \ddots & & & \ddots & \\ 0 & & a_{nn} & 0 & & b_{nn} \\ \hline c_{22} & & C'' & d_{22} & & D'' \\ & \ddots & & & \ddots & \\ 0 & & c_{nn} & 0 & & d_{nn} \end{array}\right| + (-1)^{2n+1}c_{11}b_{11}\left|\begin{array}{ccc|ccc} a_{22} & & A'' & b_{22} & & B'' \\ & \ddots & & & \ddots & \\ 0 & & a_{nn} & 0 & & b_{nn} \\ \hline c_{22} & & C'' & d_{22} & & D'' \\ & \ddots & & & \ddots & \\ 0 & & c_{nn} & 0 & & d_{nn} \end{array}\right|$$

$$= (a_{11}d_{11} - c_{11}b_{11})\left|\begin{array}{ccc|ccc} a_{22} & & A'' & b_{22} & & B'' \\ & \ddots & & & \ddots & \\ 0 & & a_{nn} & 0 & & b_{nn} \\ \hline c_{22} & & C'' & d_{22} & & D'' \\ & \ddots & & & \ddots & \\ 0 & & c_{nn} & 0 & & d_{nn} \end{array}\right|.$$

By induction this is equal to

$$(a_{11}d_{11} - c_{11}b_{11})\prod_{i=2}^{n}(a_{ii}d_{ii} - b_{ii}c_{ii}) = \prod_{i=1}^{n}(a_{ii}d_{ii} - b_{ii}c_{ii}).$$

QED

Exercise 5: Let F be a field, n a positive integer, and let V be the space of $n \times n$ matrices over F. If A is a fixed $n \times n$ matrix over F, let T_A be the linear operator on V defined by $T_A(B) = AB - BA$. Consider the family of linear operators T_A obtained by letting A vary over all diagonal matrices. Prove that the operators in that family are simultaneously diagonalizable.

Solution: If we stack the cloumns of an $n \times n$ matrix on top of each other with column one at the top, the matrix of T_A in the standard basis is then given by $\begin{bmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{bmatrix}$. Thus if A is diagonal then T_A is diagonalizable.

Now $T_AT_B(C) = ABC - ACB - BCA + CBA$ and $T_BT_A(C) = BAC - BCA - ACB + CAB$. Therefore we must show that $BAC + CAB = ABC + CBA$. The i, j-th entry of $BAC + CAB$ is $c_{ij}(a_{ii}b_{ii} + a_{jj}b_{jj})$. And this is exactly the same as the i, j-th entry of $ABC + CBA$. Thus T_A and T_B commute. Thus by Theorem 8 the family can be simultaneously diagonalized.

Section 6.6: Direct-Sum Decompositions

Exercise 1: Let V be a finite-dimensional vector space and let W_1 be any subspace of V. Prove that there is a subspace W_2 of V such that $V = W_1 \oplus W_2$.

Solution:

Exercise 2: Let V be a finite-dimensional vector space and let $W_1, \ldots, W_k$ be subspaces of V such that

$$V = W_1 + \cdots + W_k \quad \text{and} \quad \dim(V) = \dim(W_1) + \cdots + \dim(W_k).$$

Prove that $V = W_1 \oplus \cdots \oplus W_k$.

Solution:

Exercise 3: Find a projection E which projects $\mathbb{R}^2$ onto the subspace spanned by $(1, -1)$ along the subspace spanned by $(1, 2)$.

Solution:

Exercise 4: If E_1 and E_2 are projections onto independent subspaces, then $E_1 + E_2$ is a projection. True or false?

Solution:

Exercise 5: If E is a projection and f is a polynomial, then $f(E) = aI + bE$. What are a and b in terms of the coefficents of f?

Solution:

Exercise 6: True or false? If a diagonalizable operator has only the characteristic values 0 and 1, it is a projection.

Solution:

Exercise 7: Prove that if E is the projection on R along N, then $(I - E)$ is the projection on N along R.

Solution:

Exercise 8: Let $E_1, \ldots, E_k$ be linear operators on the space V such that $E_1 + \cdots + E_k = I$.

(a) Prove that if $E_i E_j = 0$ for $i \neq j$, then $E_i^2 = E_i$ for each i.
(b) In the case $k = 2$, prove the converse of (a). That is, if $E_1 + E_2 = I$ and $E_1^2 = E_1$, $E_2^2 = E_2$, then $E_1 E_2 = 0$.

Solution:

Exercise 9: Let V be a real vector space and E an idempotent linear operator on V, i.e., a projection. Prove that $(I + E)$ is invertible. Find $(I + E)^{-1}$.

Solution:

Exercise 10: Let F be a subfield of the complex numbers (or, a field of characteristic zero). Let V be a finite-dimensinal vector space over F. Suppose that $E_1, \ldots, E_k$ are projections of V and that $E_1 + \cdots E_k = I$. Prove that $E_i E_j = 0$ for $i \neq j$ (*Hint:* use the trace function and ask yourself what the trace of a porjection is.)

Solution:

Project 11: Let V be a vector space, let $W_1, \ldots, W_k$ be subspace of V, and let

$$V_j = W_1 + \cdots W_{j-1} + W_{j+1} + \cdots + W_k.$$

Suppose that $V = W_1 \oplus \cdots \oplus W_k$. Prove that the dual space V^* has the direct-sum decomposition $V^* = V_1^0 \oplus \cdots \oplus V_k^0$.

Solution: