# Linear Transformations and Matrix Algebra

Text: 1.7-2.2 Section Overviews: 1.7-2.2

| Quote of Homework Three Solutions | |
|---|---|
| **Suyuan**: This feather may look worthless, but it comes from afar and carries with it all my good intentions. | |
| | Amy Tan : The Joy Luck Club (1989) |

## 1. Review and Warm Ups

1.1. **Parameterized Vector Form.** Given the following non-homogeneous linear system,

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 4 \\ x_1 + 4x_2 - 8x_3 &= 7 \\ -3x_1 - 7x_2 + 9x_3 &= -6. \end{aligned}$$

Describe the solution set of the previous system in parametric vector form, and provide a geometric comparison with the solution to the corresponding homogeneous system.

Row reducing the augmented matrix gives,

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{rrr|r} 1 & 3 & -5 & 4 \\ 1 & 4 & -8 & 7 \\ -3 & -7 & 9 & -6 \end{array}\right] \tag{1}$$

$$\sim \left[\begin{array}{rrr|r} 1 & 0 & 4 & -5 \\ 0 & 1 & -3 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right], \tag{2}$$

which implies a general solution of the form,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix} + \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix} = \mathbf{x}_h + \mathbf{x}_p, \quad x_3 \in \mathbb{R}, \tag{3}$$

where $\mathbf{x}_h$ is the solution to the corresponding homogeneous problem and $\mathbf{x}_p$ is a particular solution for the coefficient data with this specific right-hand side. From this form we conclude the solution this this problem is a line defined by $\mathbf{x}_h$ shifted in space by the particular vector $\mathbf{x}_p$.

1.2. **Spanning Sets.** Given,

$$\mathbf{A} = \begin{bmatrix} -4 & -3 & 0 \\ 0 & -1 & 4 \\ 1 & 0 & 3 \\ 5 & 4 & 6 \end{bmatrix}. \tag{4}$$

Do the columns of $\mathbf{A}$ form a linearly independent set? What is the spanning set of the columns of $\mathbf{A}$?

Row reducing $\mathbf{A}$ to echelon form yields:

$$\begin{bmatrix} -4 & -3 & 0 \\ 0 & -1 & 4 \\ 1 & 0 & 3 \\ 5 & 4 & 6 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 4 \\ -4 & -3 & 0 \\ 5 & 4 & 6 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 4 \\ 0 & -3 & 12 \\ 0 & 4 & -9 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 4 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{bmatrix}$$

Since there are no free variables, these vectors are linearly independent.

Since there is not a pivot in every row, $A$ does not span $\mathbb{R}^4$, but since there is a pivot in each column the vectors are linearly independent. The span of this set of vectors is a three-dimensional subset of $\mathbb{R}^4$.

1.3. **Range of a Matrix Transformation.** Given,

$$\mathbf{A} = \begin{bmatrix} 1 & -3 & -4 \\ -3 & 2 & 6 \\ 5 & -1 & -8 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{5}$$

Show that there does not exist a solution to $\mathbf{Ax} = \mathbf{b}$ for every $\mathbf{b} \in \mathbb{R}^3$ and describe the set of all $\{b_1, b_2, b_3\}$ for which $\mathbf{Ax} = \mathbf{b}$ *does* have a solution.

Row reducing the corresponding augmented matrix yields:

$$\begin{bmatrix} 1 & -3 & -4 & b_1 \\ -3 & 2 & 6 & b_2 \\ 5 & -1 & -8 & b_3 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -3 & -4 & b_1 \\ 0 & -7 & -6 & b_2 + 3b_1 \\ 0 & 14 & 12 & b_3 - 5b_1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -3 & -4 & b_1 \\ 0 & -7 & -6 & b_2 + 3b_1 \\ 0 & 0 & 0 & b_1 + 2b_2 + b_3 \end{bmatrix}$$

If $b_1 + 2b_2 + b_3 \neq 0$ the system is inconsistent. However, if $b_1 + 2b_2 + b_3 = 0 \Rightarrow b_1 = -2b_2 - b_3$, the system will be consistent. Thus, for all $\mathbf{b} = b_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ the system is consistent. This forms a plane in $\mathbb{R}^3$, which represents the points in space that can be chosen for the right-hand side so that the system is soluble.

## 2. Continued Work with Language

Answer each true/false question in the chapter 1 supplemental section on page 102. It is not necessary to supply justifications but if you want your logic checked then feel free to provide a justification.

$$\{a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s,t,u,v,w,x,y,z\} = \{f,f,t,f,t,t,f,f,t,f,t,f,t,t,t,t,f,f,t,f,t,f,f,f,t,t,f\} \tag{6}$$

Please come to the office to see justifications from the solution manual.

## 3. Theory

3.1. **Characterization of a Null Mapping.** Suppose the vectors $\mathbf{v_1}, \ldots, \mathbf{v_p}$ span $\mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Suppose, as well, that $T(\mathbf{v}_i) = \mathbf{0}$ for $i = 1, \ldots, p$. Show that T is the zero-transformation. That is, show that if $\mathbf{x}$ is any vector in $\mathbb{R}^n$, then $T(\mathbf{x}) = \mathbf{0}$.

We are given that the given set of vectors spans $\mathbb{R}^n$ from this we can conclude that every vector in $\mathbb{R}^n$ can be represented as a linear combination of the given vectors. That is, for every $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x} = \sum_{i=1}^{p} c_i \mathbf{v}_i. \tag{7}$$

If we consider the linear transformation applied to this vector and note that $T(\mathbf{v}_i) = 0$ for $i = 1, 2, 3, \ldots, p$ we have,

$$T(\mathbf{x}) = T\left(\sum_{i=1}^{p} c_i \mathbf{v}_i\right) \tag{8}$$

$$= \sum_{i=1}^{p} c_i T(\mathbf{v}_i) = \mathbf{0}. \tag{9}$$

Thus, every vector in $\mathbb{R}^n$ is transformed to the zero-vector. A transformation that has this property is called a null-mapping.

3.2. **Underdetermined and Square Matrix Transformations.** If a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ then can you give a relation between $m$ and $n$? What if $T$ is also a one-to-one transformation?

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is such that it is onto $\mathbb{R}^m$ then its matrix representation is such that $T(\mathbf{x}) = \mathbf{b}$ has a solution $\mathbf{x} \in \mathbb{R}^n$ for each $\mathbf{b} \in \mathbb{R}^m$. By theorem 1.4.4 on page 43 we have that the matrix representation of $T$ has a pivot in every row. This cannot be guaranteed if there are more rows than columns. Consequently, we conclude that $m \leq n$.

If we also have that this matrix is one-to-one then the matrix representation must be such that $T(\mathbf{x}) = \mathbf{b}$ has a one solution $\mathbf{x} \in \mathbb{R}^n$ for each $\mathbf{b} \in \mathbb{R}^m$. Thus, there must be a pivot in every row and no free-variables. That is, there must also be a pivot in every column. Consequently, we conclude that $m = n$, for the mapping to also be one-to-one.

3.3. **Transforming Linear Combinations.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Show that if $T$ maps two linearly independent vectors onto a linearly dependent set of vectors, then the equation $T(\mathbf{x}) = \mathbf{0}$ has a nontrivial solution.[1]

Note that, since they are linearly independent, neither $\mathbf{u}$ nor $\mathbf{v}$ can be zero by theorem 1.7.9 on page 69. Moreover, since their linear transformations are dependent we conclude that the linear combination, $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) = \mathbf{0}$ has a non-trivial solution. That is, there exist $c_1 \neq 0$ or $c_2 \neq 0$ such that the dependence relation is satisfied. Thus, the linear transformation $T$ of their linear combination gives,

$$T(c_1\mathbf{u} + c_2\mathbf{v}) = c_1T(\mathbf{u})c_2T(\mathbf{v}) = \mathbf{0}. \tag{10}$$

However, since neither $\mathbf{u}$ nor $\mathbf{v}$ is the zero-vector and since at least one of the constants is non-zero, the combination, $c_1\mathbf{u}+c_2\mathbf{v}$, is generally non-zero and thus $T$ has a non-trivial solution.

## 4. Rotation Transformations in $\mathbb{R}^2$ and $\mathbb{R}^3$

Given,

$$\mathbf{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

4.1. **Surjective Mapping.** Show that this transformation is onto $\mathbb{R}^2$. [2]

To show that this matrix is onto $\mathbb{R}^2$ one must show that there is a pivot in every row. Though this can be done via row-reduction it will require scalings by sines/cosines and since scalings cannot be zero this restricts the values of $\theta$. I will not do this. Instead, I will argue by dependence relations. Specifically, we have page 67, a special case of theorem 1.7.7 on page 68, that says if a vector is a scaling of another then the set of vectors is linearly dependent. If this is not true then the vectors are independent and if they are independent then there is a pivot in every column of the corresponding matrix. Since this matrix is square there must also be a pivot in every row. To this end we ask if there exists a scalar $c$ such that,

$$\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = c \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}. \tag{11}$$

Suppose that $c = -\cos(\theta)/\sin(\theta)$ then the first elements would be equal. However, this now imposes the condition,

$$\sin^2(\theta) + \cos^2(\theta) = 0, \tag{12}$$

on the second variable, which is false. Thus, the first vector is not a scaling of the second and by theorem 1.7.7 they form a linearly independent set. From this it follows that there is a pivot in every row and column of $\mathbf{A}$, which implies that $\mathbf{A}$ is a one-to-one and onto transformation.

4.2. **Injective Mapping.** Show that the transformation is one-to-one. [3]

See the previous problem.

4.3. **The Unit Circle.** Show that the transformation $\mathbf{A}\hat{\mathbf{i}}$ rotates $\hat{\mathbf{i}} = [1\ 0]^{\mathrm{T}}$ counter-clockwise by an angle $\theta$ and defines a parametrization of the *unit circle*. What matrix would undo this transformation?
The affect of the transformation applied to $\hat{\mathbf{i}}$ is given by,

$$\mathbf{A}\hat{\mathbf{i}} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix},$$

which is a counterclockwise parameterization of the unit-circle for $\theta \in [0, 2\pi)$. To undo this consider the matrix $\mathbf{A}^{\mathrm{T}}$ to get $\mathbf{A}^{\mathrm{T}}\hat{\mathbf{i}} = [\cos(\theta)\ -\sin(\theta)]^{\mathrm{T}}$. This leads us to conclude that $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$, which means that $\mathbf{A}$ is a orthogonal transformation or a ridged change of coordinates.

4.4. **Determinant.** Show that $\det(\mathbf{A}) = 1$.[4]
$\det(\mathbf{A}) = \cos(\theta)\cos(\theta) - -\sin(\theta)\sin(\theta) = 1$

---

[1]Hint: Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are linearly independent but $T(\mathbf{u}), T(\mathbf{v}) \in \mathbb{R}^m$ are not. First, what must be true of $\mathbf{u}+\mathbf{v}$? Also, what must be true of $c_1, c_2$ for $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) = \mathbf{0}$? Using these facts show that $T(\mathbf{x}) = \mathbf{0}$ has a nontrivial solution.

[2]Recall that a transformation is onto if there exists an $\mathbf{x}$ for every $\mathbf{b}$ in the co-domain.

[3]Recall that a transformation is one-to-one if and only if $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.

[4]We haven't discussed determinants yet but for $2 \times 2$ there is an easy formula given by,

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies \det(\mathbf{A}) = ad - bc \tag{13}$$

, which should look familiar from homework 2.1.2 associated with requirement for having only the trivial solution.

4.5. **Orthogonality.** Show that $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$.

To formally show the orthogonality we verify either of the previous multiplications.

$$\mathbf{A}\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & \cos(\theta)\sin(\theta) - \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) - \cos(\theta)\sin(\theta) & \cos^2(\theta) + \sin^2(\theta) \end{bmatrix} = \mathbf{I}$$

4.6. **Classical Result.** Let $\mathbf{A}(\theta)\mathbf{x} = \mathbf{b}$ for each $\theta \in S$. Calculate, $\dfrac{\mathbf{x}\cdot\mathbf{b}}{|\mathbf{x}||\mathbf{b}|}$.[5] How is this related to $\theta$? [6]

This is a formula from vector calculus that should relate the dot-product of vectors to the angle between them. This result is general but in this case we have a direct comparison. That is, we have already concluded that the affect of the matrix $\mathbf{A}$ on vectors from $\mathbb{R}^2$ is to rotate them $\theta$-radians counter-clockwise in the plane. So, we already know that the angle between $\mathbf{x}$ and $\mathbf{b}$ should be $\theta$. The following derivation,

$$\frac{\mathbf{x}\cdot\mathbf{b}}{|\mathbf{x}|\,|\mathbf{b}|} = \frac{\cos(\theta)x_1^2 - \sin(\theta)x_1x_2 + \sin(\theta)x_1x_2 + \cos(\theta)x_2^2}{\sqrt{x_1^2+x_2^2}\sqrt{\cos^2(\theta)x_1^2 - \cos(\theta)\sin(\theta)x_1x_2 + \sin^2(\theta)x_2^2 + \cos^2(\theta)x_2^2 - \cos(\theta)\sin(\theta)x_1x_2 + \sin^2(\theta)x_1^2}} \tag{14}$$

$$= \cos(\theta), \tag{15}$$

shows that this fact corresponds to the standard result from calculus.

4.7. **Differentiation of Matrix Functions.** If we define the derivative of a matrix function as a matrix of derivatives then a typical product rule results. That is, if $\mathbf{A}, \mathbf{B}$ have elements, which are functions of the variable $\theta$ then $\dfrac{d}{d\theta}[\mathbf{AB}] = \mathbf{A}\dfrac{d\mathbf{B}}{d\theta} + \dfrac{d\mathbf{A}}{d\theta}\mathbf{B}$. [7] Using the identity $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, show that $\dfrac{d\left[\mathbf{A}^{-1}\right]}{d\theta} = -\mathbf{A}^{-1}\dfrac{d\mathbf{A}}{d\theta}\mathbf{A}^{-1}$. Verify this formula using the $\mathbf{A}$ matrix given above.

Assuming that the matrix $\mathbf{A}$ is a function of the variable $\theta$ and differentiating the identity $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ with respect to $\theta$ gives,

$$\frac{d}{d\theta}\left[\mathbf{A}\mathbf{A}^{-1}\right] = \frac{d}{d\theta}[\mathbf{A}]\,\mathbf{A}^{-1} + \mathbf{A}\frac{d}{d\theta}\left[\mathbf{A}^{-1}\right] \tag{16}$$

$$= \frac{d}{d\theta}[\mathbf{I}] = 0. \tag{17}$$

Solving for $\dfrac{d}{d\theta}\left[\mathbf{A}^{-1}\right]$ shows the necessary equality. Using the given matrix we have that,

$$\frac{d}{d\theta}\left[\mathbf{A}^{-1}\right] = \frac{d}{d\theta}\left[\mathbf{A}^{\mathrm{T}}\right] = \begin{bmatrix} -\sin(\theta) & \cos(\theta) \\ -\cos(\theta) & -\sin(\theta) \end{bmatrix}, \tag{18}$$

while

$$-\mathbf{A}^{-1}\frac{d\mathbf{A}}{d\theta}\mathbf{A}^{-1} = -\mathbf{A}^{\mathrm{T}}\frac{d\mathbf{A}}{d\theta}\mathbf{A}^{\mathrm{T}} \tag{19}$$

$$= -\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} -\sin(\theta) & -\cos(\theta) \\ \cos(\theta) & -\sin(\theta) \end{bmatrix}\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{20}$$

$$= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{21}$$

$$= = \begin{bmatrix} -\sin(\theta) & \cos(\theta) \\ -\cos(\theta) & -\sin(\theta) \end{bmatrix}, \tag{22}$$

which agrees with our formula.

---

[5]Recall that $\mathbf{x}\cdot\mathbf{y}$ and $|\mathbf{x}|$ are the standard dot-product and magnitude, respectively, from vector-calculus. These operations hold for vectors in $\mathbb{R}^n$ but now have the following definitions, $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{y}$ and $|\mathbf{x} = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}}$.

[6]What we are trying to extract here is the standard result from calculus, which relates the dot-product or inner-product on vectors to the angle between them. This is clear when we have vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ since we have tools from trigonometry and geometry but when treating vectors in $\mathbb{R}^n, n \geq 4$ these tools are no longer available. However, we would still like to have similar results to those of $\mathbb{R}^n, n = 2, 3$. To make a long story short, we will have these results for arbitrary vectors in $\mathbb{R}^n$ but not immediately. The first thing we must do is show that $|\mathbf{x}\cdot\mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$, which is known as Schwarz's inequality. Without this we cannot be permitted to always relate $\dfrac{\mathbf{x}\cdot\mathbf{b}}{|\mathbf{x}||\mathbf{b}|}$ to $\theta$ via inverse trigonometric functions. These details will occur in chapter 6 where we find that by using the inner-product on vectors from $\mathbb{R}^n$ we will define the notion of angle and from that distance. Using these definitions and Schwarz's inequality will then give us a triangle-inequality for arbitrary finite-dimensional vectors. This is to say that the algebra of vectors in $\mathbb{R}^n$ carries its own definition of angle and length - very nice of it don't you think? Also, it should be noted that these results exist for certain so-called infinite-dimensional spaces but are harder to prove and that the study of linear transformation of such spaces is the general setting for quantum mechanics - see MATH503:Functional Analysis for more details.

[7]To see why this is true differentiate an arbitrary element of $\mathbf{AB}$ to find $\dfrac{d}{d\theta}[\mathbf{AB}]_{ij} = \dfrac{d}{d\theta}\sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n}\dfrac{da_{ik}}{d\theta}b_{kj} + a_{ik}\dfrac{db_{kj}}{d\theta}$

4.8. **Rotations in $\mathbb{R}^3$.** Given,

$$\mathbf{R}_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \qquad \mathbf{R}_2(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \qquad \mathbf{R}_3(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Describe the transformations defined by each of these matrices on vectors in $\mathbb{R}^3$.

It is best to think about this in terms of linear combinations of columns. Consider,

$$\mathbf{R}_1(\theta)\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ \cos(\theta) \\ \sin(\theta) \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -\sin(\theta) \\ \cos(\theta) \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{A} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \end{bmatrix},$$

which implies that $\mathbf{R}_1$ leaves the $x_1$ component of the vector unchanged and rotates the $x_2, x_3$ component of the vector in the $x_2, x_3-$plane. Similar arguments show that $\mathbf{R}_2$ leaves $x_2$ unchanged and rotates the vector in the $x_1, x_3-$plane while $\mathbf{R}_3$ leave $x_3$ and rotates in the $x_1, x_2-$plane. This lays the ground-work for the so-called *Euler angles*, which provides a systematic way to rotates geometries in $\mathbb{R}^3$. An interesting consequence of the relationship between rotations in $\mathbb{R}^3$ and matrix algebra is that since these matrices do not commute, $[\mathbf{R}_1, \mathbf{R}_2] \neq \mathbf{0}$ for $\theta \neq 0$, the order one conducts the rotations matter. [8]

With this idea one can argue that $R_2$ is a counter-clockwise rotation of the $(x_1, x_3)$-plane about $x_2$ and $R_3$ is a counter-clockwise rotation of the $(x_1, x_2)$-plane about $x_3$.

## 5. Introduction to Linear Algebra for Quantum Mechanics

Define the *commutator* and *anti-commutator* of two square matrices to be,

$$[\cdot,\cdot] : \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}, \text{ such that } [\mathbf{A},\mathbf{B}] = \mathbf{AB} - \mathbf{BA}, \quad \text{for all } \mathbf{A},\mathbf{B} \in \mathbb{C}^{n\times n},$$

$$\{\cdot,\cdot\} : \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}, \text{ such that } \{\mathbf{A},\mathbf{B}\} = \mathbf{AB} + \mathbf{BA}, \quad \text{for all } \mathbf{A},\mathbf{B} \in \mathbb{C}^{n\times n},$$

respectively. Also define the *Kronecker delta* and *Levi-Civita* symbols to be,

$$\delta_{ij} : \mathbb{N} \times \mathbb{N} \to \{0,1\}, \text{ such that } \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases}$$

$$\epsilon_{ijk} : (i,j,k) \to \{-1,0,1\}, \text{ such that } \epsilon_{ijk} = \begin{cases} 1, & \text{if } (i,j,k) \text{ is } (1,2,3), (2,3,1) \text{ or } (3,1,2), \\ -1, & \text{if } (i,j,k) \text{ is } (3,2,1), (1,3,2) \text{ or } (2,1,3), \\ 0, & \text{if } i = j \text{ or } j = k \text{ or } k = i \end{cases}$$

respectively. Also define the so-called *Pauli spin-matrices* (PSM) to be,

$$\sigma_1 = \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

5.1. **The PSM are *self-adjoint* matrices.** Show that $\sigma_m = \sigma_m^{\mathrm{H}}$ for $m = 1, 2, 3$.

A matrix is self-adjoint if it is equal to its own complex conjugated transpose. That is, $\mathbf{A}$ is self-adjoint (also called Hermitian) if $\mathbf{A} = \mathbf{A}^{\mathrm{H}} = \bar{\mathbf{A}}^{\mathrm{T}}$. Notice that if a matrix has only real entries then self-ajoint implies symmetric. Clearly, for $m = 1$ and $m = 3$ the matrix is real and symmetric an therefore self-adjoint. When $m = 2$ we write,

$$\begin{aligned} \sigma_y^{\mathrm{H}} &= \bar{\sigma_y}^{\mathrm{T}} \\ &= \overline{\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}}^{\mathrm{T}} \\ &= \overline{\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}} \\ &= \begin{bmatrix} 0 & \bar{i} \\ \bar{i} & 0 \end{bmatrix} \\ &= \sigma_y. \end{aligned}$$

Hence $\sigma_y$ is self-adjoint.

[8]To see this take a textbook and orient it so that the spine is facing you and the cover is facing up. Rotate the text clockwise $\pi/2$ about the z-axis then rotate it $\pi/2$ about the y-axis. At this point I am looking at the back of the textbook and the text is upside-down. Now do the rotations in reverse order and you will see the lack of Commutativity.

5.2. **The PSM are *unitary* matrices.** Show that $\sigma_m \sigma_m^{\mathrm{H}} = \mathbf{I}$ for $m = 1, 2, 3$ where $[\mathbf{I}]_{ij} = \delta_{ij}$.
A matrix is unitary if $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}$. That is, a matrix is unitary if its adjoint is also its own inverse. Since we already know that the PSM are self-adjoint, $\sigma_m^{\mathrm{H}} = \sigma_m$, we need only show that $\sigma_m^2 = \mathbf{I}$ for $m = 1, 2, 3$.

$$
\begin{aligned}
\sigma_1^2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
& \begin{bmatrix} 0 \cdot 0 + 1 \cdot 1 & 0 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 0 + 0 \cdot 1 & 1 \cdot 1 + 0 \cdot 0 \end{bmatrix} \\
&= \mathbf{I} \\
\sigma_2^2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
&= \mathbf{I} \\
\sigma_3^2 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
&= \mathbf{I}
\end{aligned}
$$

Hence, the PSM are unitary matrices.

5.3. **Trace and Determinant.** Show that $\mathrm{tr}(\sigma_m) = 0$ and $\det(\sigma_m) = -1$ for $m = 1, 2, 3$.
Given a matrix,

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

we define the trace and determinant of $\mathbf{A}$ with the following matrix functions,

$$
T = \mathrm{tr}(\mathbf{A}) = a + d, \qquad D = \det(\mathbf{A}) = ad - bc.
$$

It is easy to see that the PSM are traceless matrices. That is $\mathrm{tr}(\sigma_m) = 0 + 0$ for $m = 1, 2$ and $\mathrm{tr}(\sigma) = 1 - 1 = 0$ for $m = 3$. Another quick check shows,

$$
\begin{aligned}
\det(\sigma_1) &= 0 \cdot 0 - 1 \cdot 1 = -1 \\
\det(\sigma_2) &= 0 \cdot 0 - i(-i) = -1 \\
\det(\sigma_3) &= 1 \cdot (-1) - 0 \cdot 0 = -1
\end{aligned}
$$

Generally, one can show that unitary matrices have determinant one or negative one.

5.4. **Anti-Commutation Relations.** Show that $\{\sigma_i, \sigma_j\} = 2\delta_{ij}\mathbf{I}$ for $i = 1, 2, 3$ and $j = 1, 2, 3$.
If we notice a couple of identities first then the busy-work is reduced. Specifically, note that $\{\mathbf{A}, \mathbf{B}\} = \{\mathbf{B}, \mathbf{A}\}$ and $\{\mathbf{A}, \mathbf{A}\} = 2\mathbf{A}^2$. So, by the self-adjoint and unitary properties we have, $\{\sigma_i, \sigma_i\} = 2\sigma_i^2 = 2\mathbf{I}$ for $i = 1, 2, 3$. For all other cases we expect the anti-commutator to

return a zero matrix. To show this we need only consider the following ordered pairs, $(i, j) = \{(1,2),\ (1,3),\ (2,3)\}$. Doing so gives,

$$
\begin{aligned}
\{\sigma_1, \sigma_2\} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \\
&= \mathbf{0} \\
\{\sigma_1, \sigma_3\} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\
&= \mathbf{0} \\
\{\sigma_2, \sigma_3\} &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} \\
&= \mathbf{0}
\end{aligned}
$$

Taken together these statements imply that $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$ for $i = 1,2,3$ and $j = 1,2,3$.

5.5. **Commutation Relations.** Show that $[\sigma_i, \sigma_j] = 2\sqrt{-1}\sum_{k=1}^{3}\epsilon_{ijk}\sigma_k$ for $i = 1,2,3$ and $j = 1,2,3$.

Again, to make quick work of this we notice that $[\mathbf{A}, \mathbf{B}] = -[\mathbf{B}, \mathbf{A}]$ and $[\mathbf{A}, \mathbf{A}] = \mathbf{0}$. This implies that if the subscripts are the same then $[\sigma_i, \sigma_i] = 2\sqrt{-1}\sum_{k=1}^{3}\epsilon_{iik}\sigma_k = 0$ since $\epsilon_{iik} = 0$ by definition. As before we now only need to determine the commutator relation for $(i, j) = \{(1,2),\ (1,3),\ (2,3)\}$ and note that if the subscripts are switched then a negative sign is introduced. Moreover, we can use the previous results since the commutator is just the anti-commutator with a subtraction. Thus,

$$
[\sigma_1, \sigma_2] = 2\sqrt{-1}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 2\sqrt{-1}\sigma_3
$$

$$
\begin{aligned}
[\sigma_1, \sigma_3] &= 2\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\
&= -2\sqrt{-1}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = -2\sqrt{-1}\sigma_2
\end{aligned}
$$

$$
[\sigma_2, \sigma_3] = 2\sqrt{-1}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2\sqrt{-1}\sigma_1
$$

this and noting the anti-symmetry of the commutator establishes the following pattern:

| $i$ | $j$ | $\sum_{k=1}^{3}\epsilon_{ijk}\sigma_k$ |
|---|---|---|
| 1 | 2 | $\epsilon_{121}\sigma_1 + \epsilon_{122}\sigma_2 + \epsilon_{123}\sigma_3 = \sigma_3$ |
| 1 | 3 | $\epsilon_{131}\sigma_1 + \epsilon_{132}\sigma_2 + \epsilon_{133}\sigma_3 = -\sigma_2$ |
| 2 | 3 | $\epsilon_{231}\sigma_1 + \epsilon_{232}\sigma_2 + \epsilon_{233}\sigma_3 = \sigma_1$ |

Noting that $\epsilon_{ijk} = -\epsilon_{jik}$ completes the pattern.

5.6. **Spin$-\frac{1}{2}$ Systems.** In quantum mechanics spin one-half particles, typically electrons[9] , have 'spins' characterized by the following vectors:

$$\mathbf{e}_u = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_d = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

where $\mathbf{e}_u$ represents spin-up and $\mathbf{e}_d$ represents spin-down.[10] The following matrices,

$$\mathbf{S}_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{S}_- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

are linear transformations, which act on $\mathbf{e}_u$ and $\mathbf{e}_d$.

5.6.1. *Projections of Spin$-\frac{1}{2}$ Systems.* Compute and describe the affect of the transformations, $\mathbf{S}_+(\mathbf{e}_u + \mathbf{e}_d)$, and $\mathbf{S}_-(\mathbf{e}_u + \mathbf{e}_d)$.
The transformations $\mathbf{S}_+$ and $\mathbf{S}_-$ have the following affects,

$$\mathbf{S}_+(\mathbf{e}_u + \mathbf{e}_d) = \mathbf{e}_u \tag{23}$$

$$\mathbf{S}_-(\mathbf{e}_u + \mathbf{e}_d) = \mathbf{e}_d, \tag{24}$$

which implies that each transformation chooses/selects the up/down vector when given their combination. That is, given a linear combination of 'up' and 'down' vectors, the matrix $\mathbf{S}_+$ will create a system which has only the 'up' direction, while $\mathbf{S}_-$ will create a system with only the 'down' direction.

5.6.2. *Properties of Projections.* $\mathbf{S}_+$ and $\mathbf{S}_-$ are projection transformations. Projection transformations are known to destroy information. Justify this in the case of $\mathbf{S}_+$ by showing that $\mathbf{S}_+$ is neither one-to-one nor onto $\mathbb{R}^2$.
These projections are not invertible mappings.This can be seen in both matrices by their pivot structures. Neither have as many pivots as columns and the invertible matrix theorem tells us that neither are invertible. Consequently, neither of the mappings are one-to-one nor onto.

---

[9] In general these particles are called fermions. `http://en.wikipedia.org/wiki/Spin-1/2`, `http://en.wikipedia.org/wiki/Fermions`

[10] In quantum mechanics, the concept of spin was originally considered to be the rotation of an elementary particle about its own axis and thus was considered analogous to classical angular momentum subject to quantum quantization. However, this analogue is only correct in the sense that spin obeys the same rules as quantized angular momentum. In 'reality' spin is an intrinsic property of elementary particles and it is the roll of quantum mechanics to understand how to associate quantized particles with spin to their associated background field in such a way that certain field properties/symmetries are preserved. This is studied in so-called quantum field theory. `http://www.physics.thetangentbundle.net/wiki/Quantum_mechanics/spin`, `http://en.wikipedia.org/wiki/Spin_(physics)`, `http://en.wikipedia.org/wiki/Quantum_field_theory`