MA281: Introduction to Linear Algebra

Dylan C. Beck

Contents

Chapter 1

Matrices and Vector Spaces

Often, real-world problems require us to deal with large amounts of data and information that can be most efficiently organized by rows and columns in what we will refer to as a matrix. We will soon see that matrices possess an arithmetic that yields a highly sophisticated and useful theory.

1.1 Matrices and Matrix Addition

Unless otherwise specified, we will assume throughout this chapter that m and n are positive integers. We say that a visual representation of any collection of data arranged into m rows and n columns is an $m \times n$ **array**. Each object of an $m \times n$ array A is a **component** or **element** of A. Each component of A can be uniquely identified by specifying its row and column. Explicitly, we use the symbol a_{ij} to indicate the component of A in the ith row and jth column; often, we will refer to a_{ij} as the (i, j)th **entry** of the array A. Collectively, therefore, we may view the array A as **indexed** by its objects a_{ij} for each pair of integers $1 \leq i \leq m$ and $1 \leq j \leq n$. Components of the form a_{ii} are referred to as the **diagonal** entries of A because they lie in the same row and column of A; the collection of all diagonal entries of A is called the **main diagonal** of A. We will adopt the convention that an $m \times n$ array be written using large rectangular brackets, as in the following.

Example 1.1.1. Consider the case that Alice, Bob, Carly, and Daryl play Bridge together. If Alice and Carly belong to one team and Bob and Daryl belong to the opposing team, then we may encode this information (i.e., these teams) as the two columns of the following 2×2 array T.

$$T = \begin{bmatrix} \text{Alice} & \text{Bob} \\ \text{Carly} & \text{Daryl} \end{bmatrix}$$

Observe that $t_{11} = \text{Alice}$, $t_{12} = \text{Bob}$, $t_{21} = \text{Carly}$, and $t_{22} = \text{Daryl}$. One could also just as well swap the rows and columns to display the teams as rows by constructing the following 2×2 array T^t.

$$T^t = \begin{bmatrix} \text{Alice} & \text{Carly} \\ \text{Bob} & \text{Daryl} \end{bmatrix}$$

Our principal concern throughout this course are those $m \times n$ arrays whose consisting entirely of real numbers. Under this restriction, we may refer to an $m \times n$ array as a (real) $m \times n$ **matrix**. Generally, one can define matrices consisting of elements of any ring, but we will not.

Example 1.1.2. Each real number x may be viewed as a real 1×1 matrix $\begin{bmatrix} x \end{bmatrix}$.

Example 1.1.3. Consider the scenario of Example 1.1.1. We may assign to each player a real number called a "skill value" between 0 and 100, e.g., suppose that Alice has skill value a; Bob has skill value b; Carly has skill value c; and Daryl has skill value d. Under this convention, the matrices of Example 1.1.1 yield new matrices called "skill matrices"; they are given by

$$S = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } S^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Our previous three examples dealt with **square** matrices, i.e., matrices for which the number of rows and the number of columns were the same (i.e., $m = n$); however, not all matrices are square.

Example 1.1.4. Consider the 1×5 matrix $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix}$ of the first five positive integers.

We refer to matrices with only one row as **row vectors**; likewise, matrices with only one column are called **column vectors**. We will return to the notion of a vector in our study of vector spaces. Often, we will also use the terminology (horizontal) n-**tuples** when discussing row vectors with n columns and (vertical) m-tuples when discussing column vectors with m rows.

Like we mentioned in the first paragraph of this section, an $m \times n$ matrix A is uniquely determined by the elements a_{ij} in its ith row and jth column for each pair of integers $1 \leq i \leq m$ and $1 \leq j \leq n$. For instance, the matrix of Example 1.1.4 is the unique matrix with one row whose jth column consists of the integer j for each integer $1 \leq j \leq 5$. Under this identification, we will adopt the one-line notation $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ for the $m \times n$ matrix A with a_{ij} in its ith row and jth column.

Example 1.1.5. Consider the 2×3 matrix whose ith row and jth column consists of the sum $i + j$. We may write this symbolically (in one-line notation) as $\begin{bmatrix} i + j \end{bmatrix}_{\substack{1 \leq i \leq 2 \\ 1 \leq j \leq 3}}$ and explicitly as

$$\begin{array}{c c} & \begin{array}{ccc} j = 1 & j = 2 & j = 3 \end{array} \\ \begin{array}{c} i = 1 \\ i = 2 \end{array} & \begin{bmatrix} 1 + 1 & 1 + 2 & 1 + 3 \\ 2 + 1 & 2 + 2 & 2 + 3 \end{bmatrix} \end{array} \text{ or } \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}.$$

Example 1.1.6. Given any positive integers m and n, there is one and only one matrix consisting entirely of zeros: it is the $m \times n$ **zero matrix**, and it is denoted by $O_{m \times n}$. Often, if we are dealing with the case that $m = n$, then we will simply abbreviate the $n \times n$ zero matrix $O_{n \times n}$ as O_n.

Example 1.1.7. We refer to the matrix $I_{m \times n} = \begin{bmatrix} \delta_{ij} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ as the $m \times n$ **identity matrix**, where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and} \\ 0 & \text{if } i \neq j \end{cases}$$

is the **Kronecker delta**. Put another way, the $m \times n$ identity matrix is the unique $m \times n$ matrix whose (i, j)th entry is 1 for each pair of integers $1 \leq i \leq m$ and $1 \leq j \leq n$ such that $i = j$ and whose other components are all zero. Explicitly, we have that

$$I_{2 \times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } I_{2 \times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Like with the zero matrix, we will simply abbreviate the $n \times n$ identity matrix $I_{n \times n}$ as I_n. Observe that the only nonzero components of I_n lie on its main diagonal, hence it is a **diagonal matrix**. Even more, by definition, it is the unique diagonal matrix whose nonzero entries are all one.

Example 1.1.8. Given any $m \times n$ matrix $A = \big[a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, its **matrix transpose** A^t is the $n \times m$ matrix obtained by swapping the rows and columns of A, i.e., we have that $A^t = \big[a_{ji}\big]_{\substack{1 \leq j \leq n \\ 1 \leq i \leq m}}$. Explicitly, for the matrix A defined in Example 1.1.5, we have that

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \text{ and } A^t = \begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{bmatrix}.$$

Observe that the first row of A becomes the first column of A^t (and likewise for the second row). Consequently, the transpose of any $1 \times n$ row vector is an $n \times 1$ column vector. We will also refer to A^t simply as the transpose of A; the process of computing A^t is called **transposition**.

Definition 1.1.9. We say that an $m \times n$ matrix A is **symmetric** if it holds that $A^t = A$. Observe that a matrix is symmetric only if it is square, i.e., a non-square matrix is never symmetric.

Considering that matrices encode numerical data, it is not surprising to find that they induce their own arithmetic. Using one-line notation, matrix addition can be defined as follows.

Definition 1.1.10. Given any $m \times n$ matrices $A = \big[a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ and $B = \big[b_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, the **matrix sum** of A and B is the $m \times n$ matrix $A + B = \big[a_{ij} + b_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. Put in words, the matrix sum $A + B$ is the $m \times n$ matrix whose (i, j)th entry is the sum of the (i, j)th entries of A and B.

Caution: the matrix sum is not defined for matrices with different numbers of rows or columns.

Example 1.1.11. If A is any $m \times n$ matrix, then we have that $A + O_{m \times n} = A = O_{m \times n} + A$. Consequently, we may view $O_{m \times n}$ as the **additive identity** among all $m \times n$ matrices.

Generally, if $A = \big[a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ is a real $m \times n$ matrix, then we will typically refer to any real number c as a **scalar**, and we define the **scalar multiple** of A by the scalar c as $cA = \big[ca_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. Essentially, we may view this as the sum of the matrix A with itself c times.

Example 1.1.12. Given any $m \times n$ matrix $A = \big[a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, let $-A = \big[-a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. We have that $A + (-A) = O_{m \times n} = -A + A$, and we say that $-A$ is the **additive inverse** of A.

Example 1.1.13. If $A = \big[a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ is any $m \times n$ matrix, then $A + A = \big[2a_{ij}\big]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.

Our next proposition illustrates that matrix transposition and matrix addition are compatible.

Proposition 1.1.14. *Let A and B be any $m \times n$ matrices. We have that $(A + B)^t = A^t + B^t$. Put another way, the transpose of a sum of matrices is the sum of the matrix transposes.*

Proof. By Definition 1.1.10, the (i, j)th entry of $A + B$ is the sum of the (i, j)th entry of A and the (i, j)th entry of B. By Example 1.1.8, the (i, j)th entry of $(A + B)^t$ is the (j, i)th entry of $A + B$, i.e., the sum of the (j, i)th entry of A and the (j, i)th entry of B. But by the same example, this is the sum of the (i, j)th entry of A^t and the (i, j)th entry of B^t. Ultimately, this shows that the (i, j)th entry of $(A + B)^t$ and the (i, j)th entry of $A^t + B^t$ are the same so that $(A + B)^t = A^t + B^t$. $\square$

We will prove in the first group work assignment that matrix transposition and scalar multiplication of matrices are compatible in a similar sense as the previous proposition. Combined, we will see that the compatibility of matrix transposition with matrix addition and scalar multiplication lead to an interesting family of square matrices that are called **skew-symmetric** matrices. Explicitly, an $n \times n$ matrix A is skew-symmetric if it satisfies that $A + A^t = O_n$, i.e., $A^t = -A$.

1.2 Rotation Matrices and Matrix Multiplication

Let $\mathbb{R}$ denote the set of real numbers. Recall that every point (x, y) in the Cartesian plane $\mathbb{R}^2$ can be written as $(r\cos\theta, r\sin\theta)$ for some real number r and some angle θ. Consequently, we may specify any point in the plane by writing $x = r\cos\theta$ and $y = r\sin\theta$. Rotation of the point (x, y) through another angle ϕ yields a new point defined by $x' = r\cos(\theta + \phi)$ and $y' = r\sin(\theta + \phi)$. Using the addition formulas for sine and cosine, we find that $x' = r(\cos\theta\cos\phi - \sin\theta\sin\phi)$ and $y' = r(\sin\theta\cos\phi + \sin\phi\cos\theta)$. Our objective in this section is to provide an efficient method of rotating points in the plane through a specified angle ϕ. We achieve this as follows.

We have seen in the previous section that any matrix can be transposed and any two matrices can be added together to obtain new matrices. Even more, if the number of columns of a matrix A is equal to the number of rows of a matrix B, then A and B can be multiplied together!

Definition 1.2.1. Given any $m \times n$ matrix $A = \left[a_{ij}\right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ and any $n \times r$ matrix $B = \left[a_{ij}\right]_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq r}}$ the **matrix product** of A and B is the $m \times r$ matrix AB whose (i, j)th entry is given by

$$AB_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}.$$

Put in words, the matrix product AB is the $m \times r$ matrix whose (i, j)th entry is the sum of the products of the (i, k)th entry of A and the (k, j)th entry of B for all integers $1 \leq k \leq n$.

Crucially, the order of A and B in the matrix product matters; the matrix product BA is defined analogously. Be sure to notice also that the number of rows of AB is the same as the number of rows of A, and the number of columns of AB is the same as the number of columns of B.

Caution: the product is not defined for matrices with an incompatible number of rows and columns.

Example 1.2.2. Consider the following matrices.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$$

Considering that A is a 2×3 matrix and B is a 3×2 matrix, both of the products AB and BA can be formed: AB is a 2×2 matrix, and BA is a 3×3 matrix. Explicitly, they are as follows.

$$AB = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1(-1) + 2(0) + 3(-1) & 1(0) + 2(1) + 3(1) \\ 2(-1) + 3(0) + 4(-1) & 2(0) + 3(1) + 4(1) \end{bmatrix} = \begin{bmatrix} -4 & 5 \\ -6 & 7 \end{bmatrix}$$

$$BA = \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} -1(1) + 0(2) & -1(2) + 0(3) & -1(3) + 0(4) \\ 0(1) + 1(2) & 0(2) + 1(3) & 0(3) + 1(4) \\ -1(1) + 1(2) & -1(2) + 1(3) & -1(3) + 1(4) \end{bmatrix} = \begin{bmatrix} -1 & -2 & -3 \\ 2 & 3 & 4 \\ 1 & 1 & 1 \end{bmatrix}$$

Example 1.2.3. We say that an $n \times n$ matrix A is a **Markov matrix** if each component of A is a non-negative real number and the sum of each column of A is 1. For instance, the 2×2 matrix

$$A = \begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix}$$

is a Markov matrix. We may view this Markov matrix as representing a real-life scenario as follows.

Godspeed You! Black Emperor are playing at the Blue Note in Columbia, Missouri, and Alice and Bob are considering attending the concert. Currently, Alice is 90% certain that she will attend, so she is 10% certain that she will not attend. On the other hand, Bob is only 50% sure he will attend. Consequently, the columns of the matrix A represent Alice and Bob, respectively, and the rows represent their certainty or uncertainty that they will attend the concert, respectively.

Even more, suppose that today, Alice has the propensity a to attend the concert and Bob has the propensity b to attend, and tomorrow, Alice has the propensity $0.9a + 0.5b$ to attend the concert and Bob has the propensity $0.1a + 0.5b$ to attend. If $p = \begin{bmatrix} a & b \end{bmatrix}^t$ is the "propensity vector," then tomorrow, the propensity that Alice and Bob will attend the concert is given by the matrix product

$$Ap = \begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.9a + 0.5b \\ 0.1a + 0.5b \end{bmatrix}.$$

We could continue to iterate this process to predict the propensity that Alice and Bob will attend the concert on any given day in the future; this is called a **Markov process**.

We will demonstrate now that matrix multiplication is associative and distributive.

Proposition 1.2.4. *If A is any $m \times n$ matrix, B is any $n \times r$ matrix, and C is any $r \times s$ matrix, then the matrix products $A(BC)$ and $(AB)C$ are well-defined; in fact, they are equal.*

Proof. By Definition 1.2.1, we have that BC is an $n \times s$ matrix, hence the matrix product $A(BC)$ is well-defined because the number of columns of A is equal to the number of rows of BC; a similar argument shows that $(AB)C$ is well-defined, hence it suffices to prove that $A(BC) = (AB)C$. By the same definition, the (i, j)th entry of $A(BC)$ is the sum of the products of the (i, k)th entry of A and the (k, j)th entry of BC for all integers $1 \leq k \leq n$, and the (k, j)th entry of BC is the sum of the products of the (k, ℓ)th entry of B and the (ℓ, j)th entry of C for all integers $1 \leq \ell \leq r$. Put into symbols, the previous sentence can be expressed as the double summation identity

$$A(BC)_{ij} = \sum_{k=1}^{n} \sum_{\ell=1}^{r} a_{ik} b_{k\ell} c_{\ell j}.$$

Considering that the order of summation of a finite sum does not matter, it follows that

$$A(BC)_{ij} = \sum_{\ell=1}^{r} \sum_{k=1}^{n} a_{ik} b_{k\ell} c_{\ell j}.$$

Observe that $\sum_{k=1}^{n} a_{ik} b_{k\ell}$ is nothing more than the (i, ℓ)th entry of AB, hence we may view the (i, j)th entry of $A(BC)$ as the sum of the products of the (i, ℓ)th entry of AB and the (ℓ, j)th entry of C for all integers $1 \leq i \leq r$, i.e., it is the (i, j)th entry of $(AB)C$. Ultimately, this shows that the (i, j)th entry of $A(BC)$ and the (i, j)th entry of $(AB)C$ are the same so that $A(BC) = (AB)C$. $\square$

Example 1.2.5. If A is any $n \times n$ matrix, then the matrix product of A with itself is denoted simply by A^2; it is itself an $n \times n$ matrix, hence we may form the matrix product of A^2 with A. By Proposition 1.2.4, the matrices $(A^2)A$ and $A(A^2)$ are equal; they are denoted simply by A^3. Continuing in this manner, the k-fold product of A is $A^k = AA^{k-1} = A^{k-1}A$ for all integers $k \geq 2$.

Proposition 1.2.6. *If A is any $m \times n$ matrix and B and C are any $n \times r$ matrices, then the product $A(B+C)$ is well-defined; $A(B+C) = AB + AC$; and $A(cB) = c(AB)$ for all scalars c.*

Proof. By Definition 1.1.10, the matrix sum $B+C$ is an $n \times r$ matrix, hence the product $A(B+C)$ is well-defined because the number of columns of A is equal to the number of rows of $B+C$. By Definition 1.2.1, the (i,j)th entry of $A(B+C)$ is the sum of the products of the (i,k)th entry of A and the $(k,j)$the entry of BC for all integers $1 \leq k \leq n$; the latter is by Definition 1.1.10 the sum of the (k,j)th entry of B and the (k,j)th entry of C. Because addition is multiplication is distributive, the (i,j)th entry of $A(B+C)$ is the sum of the products of the (i,k)th entry of A and the (k,j)th entry of B for all integers $1 \leq k \leq n$ plus the sum of the products of the (i,k)th entry of A and the (k,j)th entry of C for all integers $1 \leq k \leq n$, i.e., it is the sum of the (i,j)th entry of AB and the (i,j)th entry of AC, i.e., it is the (i,j)th entry of $AB + AC$. Because the (i,j)th entry of $A(B+C)$ and the (i,j)th entry of $AB + AC$ are the same, we conclude that $A(B+C) = AB + AC$.

We leave it as an exercise for the reader to demonstrate that $A(cB) = c(AB)$ for all scalars c; however, we remark that the proof is similar to the proof of Proposition 1.2.4. □

Ultimately, Proposition 1.2.6 implies that matrix multiplication is distributive, i.e., if A is any $m \times n$ matrix, B and C are any $n \times r$ matrices, and c is any scalar, then $A(cB+C) = c(AB) + AC$. Even more, like matrix addition, matrix multiplication is compatible with transposition.

Proposition 1.2.7. *If A is any $m \times n$ matrix and B is any $n \times r$ matrix, then $(AB)^t = B^t A^t$. Put another way, the transpose of a matrix product is the reverse matrix product of the transposes.*

Proof. By Example 1.1.8, the (i,j)th entry of $(AB)^t$ is the (j,i)th AB. By Definition 1.2.1, the (j,i)th entry of AB is the sum of the products of the (j,k)th entry of A and the (k,i)th entry of B for all integers $1 \leq k \leq n$. Considering that scalar multiplication is commutative, this is equal to the sum of the products of the (i,k)th entry of B^t and the (k,j)th entry of A^t for all integers $1 \leq k \leq n$, i.e., it is the (i,j)th entry of $B^t A^t$. We conclude therefore that $(AB)^t = B^t A^t$. □

We return now to the setup of the first paragraph of this section. Once again, we are considering some point (x,y) in the Cartesian plane, and we are identifying this point by its polar coordinates $x = r\cos\theta$ and $y = r\sin\theta$ for some real number r and some angle θ. Our aim is to efficiently write down the rotation of (x,y) through another angle ϕ, resulting in a new point determined by $x' = r\cos(\theta+\phi)$ and $y' = r\sin(\theta+\phi)$. By the addition formulas for sine and cosine, it follows that $x' = r(\cos\theta\cos\phi - \sin\theta\sin\phi)$ and $y' = r(\sin\theta\cos\phi + \sin\phi\cos\theta)$. Consider the matrices

$$R(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \text{ and } X(r,\theta) = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}.$$

Observe that $X(r,\theta)$ is the column vector corresponding to the point (x,y) in the Cartesian plane, i.e., it encodes the same data as the point (x,y). By Definition 1.2.1, we have that

$$R(\phi)X(r,\theta) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} = \begin{bmatrix} r(\cos\theta\cos\phi - \sin\theta\sin\phi) \\ r(\sin\phi\cos\theta + \sin\theta\cos\phi) \end{bmatrix} = \begin{bmatrix} r\cos(\theta+\phi) \\ r\sin(\theta+\phi) \end{bmatrix}.$$

Considering that the last matrix in the above displayed equation is exactly equal to the column vector $X(r, \theta + \phi)$, i.e., the column vector corresponding to the point (x', y'), we conclude that the multiplication by the matrix $R(\phi)$ has the effect of rotating the point (x, y) in the Cartesian plane through the angle ϕ. Consequently, we refer to the matrix $R(\phi)$ as a **rotation matrix**.

Example 1.2.8. Consider the point $(1, 0)$ in the Cartesian plane. Observe that in polar coordinates, this point is determined by $r \cos \theta = 1$ and $r \sin \theta = 0$, hence we obtain the column vector

$$X(r, \theta) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

By the previous paragraph, to rotate $X(r, \theta)$ through the angle $\phi = \pi/4$, we multiply by the matrix

$$R(\pi/4) = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}.$$

Consequently, we find that rotating the point $(1, 0)$ through the angle $\phi = \pi/4$ results in the point

$$X(r, \theta + \phi) = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}.$$

But if we consider the fact that the point $(1, 0)$ lies on the unit circle and corresponds to the angle $\theta = 0$, then the point obtained by rotating $(1, 0)$ through the angle of $\phi = \pi/4$ must be exactly the point on the unit circle corresponding to the angle $\pi/4$, i.e., it must be $(\sqrt{2}/2, \sqrt{2}/2)$.

1.3 Chapter 1 Overview

Blah, blah, blah.