※ Foundations 1: Probability and Complex Numbers (1/8)

Humor is the ability to see three sides to one coin - Ned Rorem

1.1 A single coin

You may be familiar with the concept of a **sample space** from probability. This is simply the collection of possible outcomes given a probabilistic procedure. In this course, we will often discuss **state spaces** instead, which are the set of possible states some system can be in. The distinction is that a state represents a dynamic system, where the state is expected to change as time progresses due to some external force applied to it. In our case, this force will be modeled using computational gates, but more on that later.

Suppose we have a biased coin that lands on heads with probability p.

Question 1. What are the possible states that the coin can be in?

An important thing to note is that when we say "state", this corresponds to our best representation of the object, not necessarily the true physical state of the object. Think of it as a mental model for our prediction of the state of the coin. Critically, this means that our mental representation of the coin can change depending on whether we are *looking* at the coin or not.

Question 2. Suppose we took the coin and place it in a box, close it, then give it a good shake. How can we mathematically model and represent the action of shaking the box?

The above examples were instances of some important mathematical tools we will be using.

Definition 1.1 (Probability Vector)**.** A **probability vector** is a vector containing nonnegative real entries that sum to 1. The entries store the probability of seeing the event corresponding to the index.

Definition 1.2 (Stochastic Matrix)**.** A **stochastic matrix** is a matrix with nonnegative elements whose columns add up to 1.

Question 3. Show that multiplying a probability vector by a stochastic matrix always results in another probability vector.

Question 4. Describe the stochastic matrix representing the action of turning the box upside down. Think of it as flipping the coin regardless of its face.

Question 5. Consider the following coin game using a fair coin (probability of heads is $1/2$), where the action will change depending on the state. The action during a single turn is the following:

- If the current state is HEADS: Do a fair coin flip.

- If the current state is TAILS: Turn the coin over.

Describe the stochastic matrix corresponding to this game.

   Challenge: If I had someone play this game for me for 100 turns, how would I represent the state of the coin? 1000 turns? Infinite turns? Is there a state the coin will converge to?

Example 1.3. In general, we can describe the entries of a stochastic matrix for a coin by the following:

$$
\begin{bmatrix}
\mathbb{P}(T_{\text{after}}|T_{\text{before}}) & \mathbb{P}(T_{\text{after}}|H_{\text{before}}) \\
\mathbb{P}(H_{\text{after}}|T_{\text{before}}) & \mathbb{P}(H_{\text{after}}|H_{\text{before}})
\end{bmatrix}
\tag{1}
$$

This is consistent with our definition, as the columns correspond to conditioned probability distributions implying that the entries are nonnegative and do sum to 1.

1.2   Two coins

How can we extend this mathematical model to a system of two coins? Suppose we have two biased coins, where the state of the first coin is represented by the vector $\begin{bmatrix} a \\ b \end{bmatrix}$ and the state of the second coin is represented by the vector $\begin{bmatrix} c \\ d \end{bmatrix}$. Then,

$$
\begin{bmatrix}
\mathbb{P}[TT] \\
\mathbb{P}[TH] \\
\mathbb{P}[HT] \\
\mathbb{P}[HH]
\end{bmatrix}
=
\begin{bmatrix}
ac \\
ad \\
bc \\
bd
\end{bmatrix}
=
\begin{bmatrix} a \\ b \end{bmatrix}
\otimes
\begin{bmatrix} c \\ d \end{bmatrix}
=
\begin{bmatrix}
a \begin{bmatrix} c \\ d \end{bmatrix} \\
b \begin{bmatrix} c \\ d \end{bmatrix}
\end{bmatrix}
\tag{2}
$$

In the above equation, the symbol "$\otimes$" is called the **tensor product** or Kronecker product and will be the standard way we combine two state spaces.

Question 6. Suppose someone flipped two fair coins (probability of seeing heads is 1/2). How would we represent the full system of the two coins?

Two coins introduces some complexity to our model. It turns out that not all probability vectors with four elements can be described by simply taking the tensor product between two probability vectors with two elements! We will prove this is the case by taking a special example in the following question.

Question 7. Consider the following probability vector:

$$B = \begin{bmatrix} 1/2 \\ 0 \\ 0 \\ 1/2 \end{bmatrix} \tag{3}$$

Prove that this vector cannot be constructed by taking the tensor product between two probability vectors representing coins.

Question 8. Suppose both of the coins begin in the state tails. Is there a sequence of actions on the individual coins (think stochastic matrices) such that the final state of the two coins will be B defined in equation (3)?

We will call states that can't be decomposed into a tensor product of smaller states a **correlated state**. Why correlated? If I told you what the state of the first coin, this will tell you something about the state of the second coin!

1.3   Complex Numbers and Trigonometry

The shortest path between two truths in the real domain passes through the complex domain.

- Jacques Hadamard

Before diving into our exploration of quantum states, it'll be helpful to review some important definitions and properties about complex numbers, as well as the vectors and matrices that have complex entries. We will also observe the versatility of linear algebra, as it gives us a handle on describing states *and* complex numbers.

Definition 1.4 (Complex Number (Standard Representation)). A **complex number** α is a number that can be written as

$$\alpha = a + bi \tag{4}$$

for two real numbers a and b, and i is defined to be the constant satisfying $i^2 = -1$. This is the **standard representation** (or **standard form**) of expressing a complex number. The set of all complex numbers will be written as $\mathbb{C}$.

Every complex number has a **complex conjugate**. The complex conjugate of $\alpha = a + bi$ is

$$\alpha^* = a - bi. \tag{5}$$

A single complex number can be described as a vector in the **complex plane** where the x-axis corresponds to the real component a and the y-axis corresponds to the imaginary component b. We could rewrite a complex number as a column vector using its "coordinates" as $\begin{bmatrix} a \\ b \end{bmatrix}$.
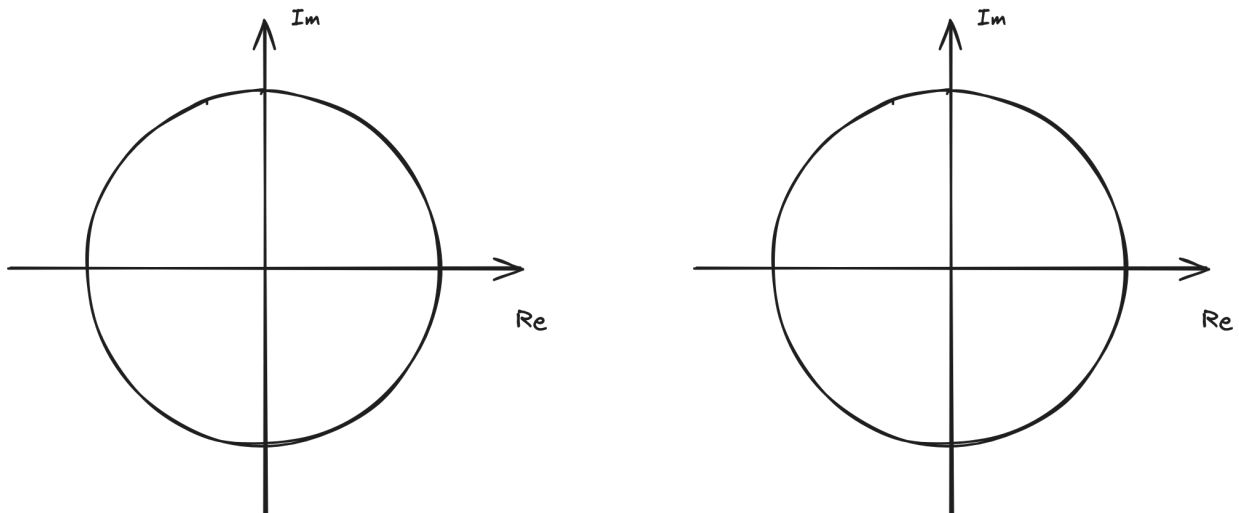
Question 9. Draw the vector representing the complex number $w = 0.5 - 0.3i$ and w^* in the left figure. In the right figure, draw $z = \sqrt{\frac{3}{4}} + i\sqrt{\frac{1}{4}}$ and z^*.

Definition 1.5 (Norm of Complex Number). The **norm** $|\alpha|$ of a complex number α is defined as

$$|\alpha| = \sqrt{a^2 + b^2}. \tag{6}$$

This is the "size" of the complex number, and we can see why this is true by looking at the complex plane. Because of this, $|\alpha| > 0$ for any complex number, and the only time $|\alpha| = 0$ is when $\alpha = 0$.

Question 10. Find the norm of w and z from the previous problem.

What can we say about a right triangle with hypotenuse of 1?
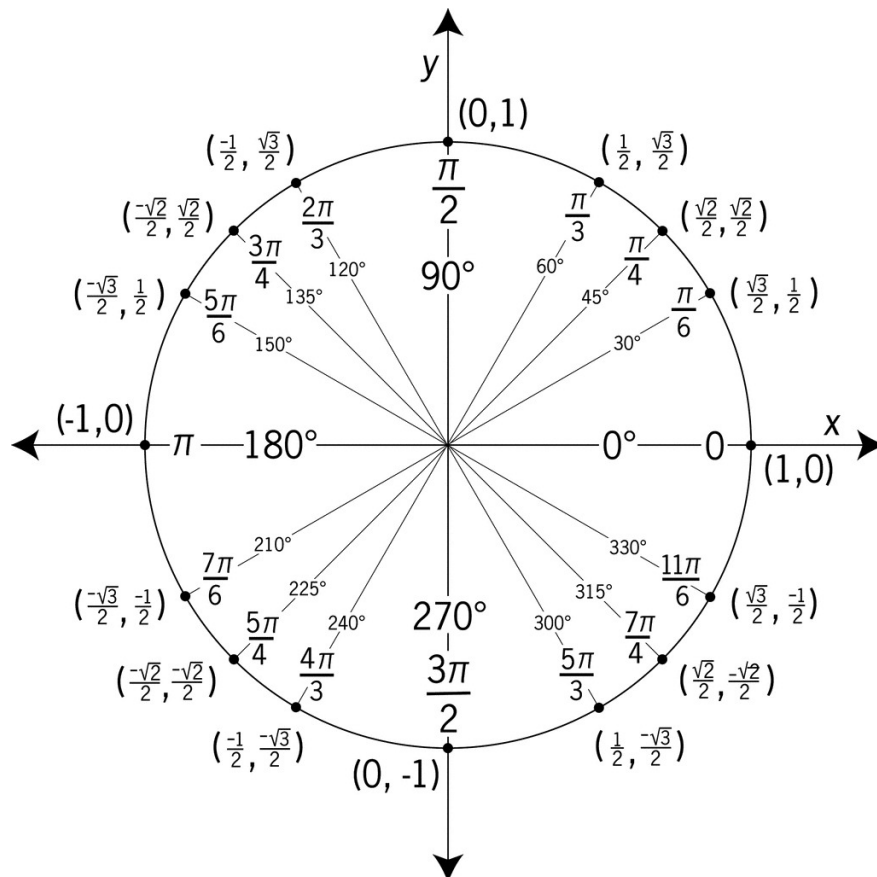
Another way to represent a complex number is by its phase representation.

Definition 1.6 (Complex Number (Phase Representation)). On the complex plane, we can also represent the number using the counterclockwise angle θ from $1 + 0i$, and its norm $|\alpha|$. That is,

$$\alpha = |\alpha|(\cos \theta + i \sin \theta) = |\alpha|e^{i\theta} \tag{7}$$

where the last equality uses the identity $e^{i\theta} = \cos \theta + i \sin \theta$.

Question 11. Write the complex number $\alpha = \frac{3}{\sqrt{2}} - \frac{3}{\sqrt{2}}i$ in its phase representation. What can you say about its complex conjugate α^*?

Question 12. Show that $|\alpha| = \sqrt{\alpha^*\alpha} = \sqrt{\alpha\alpha^*}$. Try computing the norm of α from the above example using this method.

1.4   Summary and future directions

We defined a probability vector to be a vector that models a distribution over states. We did this by assigning two quite natural constraints,

1. the entries must sum to 1, and

2. the entries must be nonnegative real numbers.

As mathematicians, we like to generalize requirements on interesting objects. How would we generalize the above definition?

The first constraint is equivalent to saying that the $L1$-norm of the vector must be 1 (we will review norms, but if this doesn't ring a bell you should do a quick google search on the definition).

Question 13. Draw the set of points in the 2D plane whose $L1$-norm is 1.

This, however, is not the standard norm we study in linear algebra! We are more familiar with measuring the length of a vector by the $L2$-norm. Maybe we can require the vectors to be unit vectors in the standard $L2$-norm.

Once we begin considering $L2$-norms, the entries don't need to be nonnegative real numbers, since we will be squaring them anyways. Maybe we can use negative numbers, or even complex numbers!

It turns out that these two generalizations for a probability vector are exactly what physicists use to describe quantum states, and more importantly what we will be using to represent the states of quantum computers. We will show that this is a good generalization as many of the questions we answered here will work for quantum states as well.

※ Homework 1 Cover Sheet

Submit this document along with your completed homework to receive credit for completing homework 1. The homework problems are on the course website. Make sure to spend some time reviewing the content before starting your one hour timer for each problem. Do not spend more than an hour per problem. If you want guidance after your hour before the solutions are released, come find us in office hours and we are more than happy to help out.

2.1 Problem 1

- Time taken:

- Self-graded score:

2.2 Problem 2

- Time taken:

- Self-graded score:

2.3 Problem 3

- Time taken:

- Self-graded score:

2.4 Problem 4

- Time taken:

- Self-graded score: