# DSC 40A Fall 2025 - Group Work Week 7

due Monday, Nov 10th at 11:59PM

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. **One person** from each group should submit your solutions to Gradescope and **tag all group members** so everyone gets credit.

This worksheet won't be graded on correctness, but rather on good-faith effort. Even if you don't solve any of the problems, you should include some explanation of what you thought about and discussed, so that you can get credit for spending time on the assignment.

In order to receive full credit, you must work in a group of two to four students for at least 50 minutes in your assigned discussion section. You can also self-organize a group and meet outside of discussion section for 80 percent credit. You may not do the groupwork alone.

## 1 Gradient Descent

Gradient descent is an algorithm used to minimize differentiable functions. In this problem, you will carry out several steps of gradient descent on a simple function of two variables.

**Problem 1.**

Consider the quadratic function

$$f(x,y) \;=\; x^2 + 3y^2 - 4x + 6y + 7.$$

**a)** Compute the gradient $\nabla f(x,y)$.

**Solution:**

$$\nabla f(x,y) = \begin{pmatrix} 2x-4 \\ 6y+6 \end{pmatrix}.$$

**b)** Write the gradient descent update rule for $f$ with step size $\alpha = 0.2$. Write your answer two ways: first using the vector form $\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix}$, and then as two separate equations for $x_{k+1}, y_{k+1}$, respectively.

**Solution:** Vector form:

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \alpha \nabla f(x_k, y_k) = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - 0.2 \begin{pmatrix} 2x_k - 4 \\ 6y_k + 6 \end{pmatrix}.$$

Scalar form:

$$x_{k+1} = x_k - 0.2(2x_k - 4) = 0.6x_k + 0.8, \qquad y_{k+1} = y_k - 0.2(6y_k + 6) = -0.2y_k - 1.2.$$

**c)** Starting from $(x_0, y_0) = (5, 4)$ and using $\alpha = 0.2$, perform *five* iterations of gradient descent. As you carry out your calculations, fill in the table below at each step.

| $k$ | $x_k$ | $y_{k+1}$ | $f(x_k, y_k)$ | $\nabla f(x_k, y_k)$ |
|---|---|---|---|---|
| 0 | | | | |
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |
| 5 | | | | |

You can and should use a calculator as you carry out these calculations.

**Solution:**

| $k$ | $x_k$ | $y_{k+1}$ | $f(x_k, y_k)$ | $\nabla f(x_k, y_k)$ |
|---|---|---|---|---|
| 0 | 5 | $-2.0$ | 84 | $\left(6,\ 30\right)$ |
| 1 | 3.8 | $-0.80$ | 6.24 | $\left(3.6,\ -6\right)$ |
| 2 | 3.08 | $-1.040$ | 1.2864 | $\left(2.16,\ 1.2\right)$ |
| 3 | 2.64 | $-0.99$ | 0.42 | $\left(1.29,\ -0.24\right)$ |
| 4 | 2.39 | $-1.00$ | 0.15 | $\left(0.77,\ 0.05\right)$ |
| 5 | 2.23 | $-0.99$ | 0.05 | $\left(0.46,\ -0.01\right)$ |

**d)** Using calculus, find the *exact* minimizer $(x^\star, y^\star)$ of $f$ and the *exact* minimum value $f(x^\star, y^\star)$.

**Solution:** Solve $\nabla f(x, y) = 0$: $2x - 4 = 0 \Rightarrow x^\star = 2$ and $6y + 6 = 0 \Rightarrow y^\star = -1$. Then

$$f(x^\star, y^\star) = (2-2)^2 + 3(-1+1)^2 = 0.$$

## 2 Probability

Most probability questions can be solved by applying one of the basic probability rules in the right way. Sometimes, some cleverness is needed to define the right sample space or the right events. There are often many ways to solve the same problem, some easier than others. It's really useful to learn multiple ways of doing the same problem, which will help you develop your problem-solving skills.

Here are the basic probability rules you'll need to use to solve the questions that follow.

**Addition Rule:**

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

**Multiplication Rule:**

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B|A)$$

**Complement Rule:**

$$\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$$

**Conditional Probability:**

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

**Problem 2.**

A bitstring is a sequence of 0s and 1s. For example, 0110100 is a bitstring of length 7.

Suppose that we generate a bitstring of length 4 such that each digit is equally likely to be a 0 or 1.

**a)** What is the probability that the bitstring is 1111?

> **Solution:** $\frac{1}{16}$.
>
> The value of one digit has no impact on the values of other digits, and each digit has a $\frac{1}{2}$ chance of being 0 or 1. Hence, the chance that all digits are 1 is $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$.

**b)** What's the probability that the bitstring contains at least one 0 and one 1?

> **Solution:** $\frac{14}{16} = \frac{7}{8}$.
>
> Note that the complement of the event "the bitstring contains at least one 0 and one 1 is "the bitstring is all 1s or the bitstring is all 0s". The latter two events are mutually exclusive, meaning they can't happen at the same time. The chance that all of the digits are 1 is $\frac{1}{16}$, as we saw in the previous part, and that's the same as the chance that all digits are 0. Hence, the chance that all digits are the same is $\frac{2}{16} = \frac{1}{8}$, and so the probability that there is at least one 0 and one 1 is $1 - \frac{1}{8} = \frac{7}{8}$.

**c)** What is the probability that a bitstring has more 0s than 1s?

**Solution:** $\frac{5}{16}$.

If the bitstring has more zeros than ones, then there are either three zeros or four zeros.

There is only one way to have four zeros: 0000.

There are 4 ways to have three zeros: 1000, 0100, 0010, 0001.

In total, the event has $1 + 4 = 5$ elements, and so its probability is $\frac{5}{16}$.

**d)** What is the probability that a bitstring has more 0s than 1s, if we know that the first bit is a 0?

**Solution:** $\frac{1}{2}$.

We can use the conditional probability formula for $P(B|A)$ with $B$ being the event that there are more 0s than 1s and $A$ being the event that the first bit is a zero. Then

$$\begin{aligned} B &= \{0000, 1000, 0100, 0010, 0001\} \\ A &= \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111\} \\ A \cap B &= \{0000, 0100, 0010, 0001\} \end{aligned}$$

Therefore

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{4}{16}}{\frac{8}{16}} = \frac{1}{2}.$$

**e)** Suppose now that you generate two bitstrings and look at one of them. You see that this bitstring has more 0s than 1s. What is the probability that in total, for both strings together, there are more 0s than 1s?

**Solution:** $\frac{59}{80}$.

We will use as our sample space ordered pairs of bitstrings, where the first bitstring in the pair represents the string that you looked at. The number of ordered pairs in $S$ is $16^2 = 256$ since there are 16 possibilities for each bitstring in the pair. We will again use the conditional probability formula for $P(B|A)$ with $B$ being the event that there are more 0s than 1s in both strings in total and $A$ being the event that the first bitstring in the pair has more zeros than ones. We know there are five bitstrings with more 0s than 1s:

$$0000, 1000, 0100, 0010, 0001$$

Therefore, since the first bitstring could have been any one of 16 strings with equal probability, we have $P(A) = \dfrac{5}{16}$.

Next, we calculate $P(A \cap B)$. If the first bitstring was 0000, then any second bitstring with at least one 0 would result in the pair having more 0s than 1s. Since 1111 is the only bitstring with no 0s, this means any of the other 15 bitstrings could be paired with 0000 to give a pair with more 0s than 1s. Otherwise, the first bitstring had three 0s and any second bitstring with at least two 0s would result in the pair having more 0s than 1s. There are eleven bitstrings that have at least two 0s. Therefore, the number of ordered pairs of bitstrings where the first bitstring has more 0s than 1s and the pair has more 0s than 1s in total is given by $1 \cdot 15 + 4 \cdot 11 = 59$. Since there are 256 pairs of strings in total, $P(A \cap B) = \dfrac{59}{256}$.

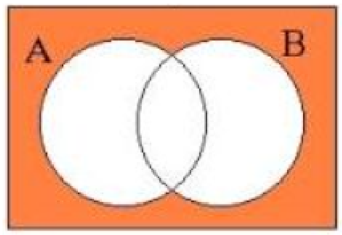
Therefore

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{59}{256}}{\frac{5}{16}} = \frac{59}{80}$$

**Problem 3.**

Let $A$ and $B$ be two independent events in the sample space $S$. Show that $\bar{A}$ and $\bar{B}$ must be independent of one another.

You may use the fact that $P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B)$, which should be apparent from the Venn diagram below.




**Solution:** Let A, B be independent so $P(A \cap B) = P(A) * P(B)$. To show that $\bar{A}$ and $\bar{B}$ are independent, we must show that $P(\bar{A} \cap \bar{B}) = P(\bar{A}) * P(\bar{B})$.

Starting with the given fact,

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= 1 - P(A \cup B) \\ &= 1 - \big(P(A) + P(B) - P(A \cap B)\big) \qquad \text{by the addition rule} \\ &= 1 - P(A) - P(B) + P(A) * P(B) \\ &= (1 - P(A)) * (1 - P(B)) \\ &= P(\bar{A}) * P(\bar{B}). \end{aligned}$$

This shows that when $A$ and $B$ are independent, so are their complements.

**Problem 4.**

Suppose you have 6 pairs of socks in your sock drawer, each in a different pattern. It is still dark out in the morning when you get dressed, so you randomly pull one sock at a time out of the drawer, until you have removed two matching socks. What is the probability that you pull out exactly 5 socks from your sock drawer in the morning?

**Solution:** $\frac{8}{33}$.

In order for you to pull out exactly 5 socks,

- the first sock can be any sock $\Rightarrow$ Prob $= 1$
- the second sock can be any of the remaining 11 socks, except for the match to the first sock $\Rightarrow$ Prob $= \frac{10}{11}$
- the third sock can be any of the remaining 10 socks, except for the matches to the first sock and the second sock $\Rightarrow$ Prob $= \frac{8}{10}$
- the fourth sock can be any of the remaining 9 socks, except for the matches to the first sock, second sock, and third sock $\Rightarrow$ Prob $= \frac{6}{9}$
- the fifth sock must match one of the first four socks, and is selected from the remaining 8 socks $\Rightarrow$ Prob $= \frac{4}{8}$

The probability of all of these things happening is

$$1 \cdot \frac{10}{11} \cdot \frac{8}{10} \cdot \frac{6}{9} \cdot \frac{4}{8} = \frac{8}{33}$$

**Problem 5.**

You're listening to a YouTube playlist of the 14 songs in Oscar Peterson's album, *Solo*. Suppose you're listening on shuffle, and each time a new song starts, it's equally likely to be any of the 14 songs on the playlist, regardless of which songs have been played so far. How many songs must you listen to so that the probability of hearing "Mirage" is at least 75%?

**Solution:** 19.

For each song, the probability that it is "Mirage" is $\dfrac{1}{14}$ and the probability that it is not "Mirage" is $\dfrac{13}{14}$. The probability that after $n$ songs you haven't heard "Mirage" is $\left(\dfrac{13}{14}\right)^n$. Using the complement rule, this means that the probability you have heard "Mirage" at least once after playing $n$ songs is $1 - \left(\dfrac{13}{14}\right)^n$. This means we're looking for the smallest integer $n$ such that $1 - \left(\dfrac{13}{14}\right)^n > \dfrac{3}{4}$. We can solve for $n$ in this inequality by taking a logarithm of both sides, as follows.

$$\begin{aligned}
1 - \left(\frac{13}{14}\right)^n &> \frac{3}{4} \\
\frac{1}{4} &> \left(\frac{13}{14}\right)^n \\
\ln\left(\frac{1}{4}\right) &> n \ln\left(\frac{13}{14}\right) \\
\frac{\ln\left(\frac{1}{4}\right)}{\ln\left(\frac{13}{14}\right)} &< n \qquad \text{inequality flips when dividing by a negative number} \\
18.7 &< n
\end{aligned}$$

So, you would have to listen to at least 19 songs to ensure that you have a 75% chance of hearing "Mirage". This is longer than the full playlist, so if you really wanted to hear this song, maybe shuffle wasn't the greatest idea.

**Problem 6.**

Suppose we scramble the 26 letters of the alphabet in a random order so that each rearrangement is equally likely. What is the probability that the letters ABC wind up next to each other in that order?

**Solution:** $\frac{1}{650}$.

Consider as the sample space all possible rearrangements of the 26 letters. There are 26 options for what goes first, then 25 options for what comes next, then 24, and so on. Therefore, the total number of rearrangements is $26! = 26 \cdot 25 \cdot \dots \cdot 2 \cdot 1$. Each rearrangement is equally likely, so we can calculate the probability of ABC ending up next to each other in that order by counting the number of rearrangements where that happens and dividing by the total number of rearrangements, 26!.

For a rearrangement to have $ABC$ consecutively, we have 24 options of which three positions out of the 26 they can fill. These three letters can be in any adjacent positions, so the $A$ can be placed in any position except for the last two positions, otherwise we wouldn't have space to fit the $B$ and $C$. Once the position of the $ABC$ is determined, we have filled 3 positions and have 23 positions empty. We can fill the first of these open spots with any of the remaining 23 letters of the alphabet, excluding $ABC$. Then for the next open spot, we have 22 options, then 21 and so on. Therefore, the number of rearrangements with $ABC$ consecutively is given by $24 \cdot 23! = 24!$. Another way to look at this is to think of rearranging 24 items, which are $ABC, D, E, F, \dots, Z$, and so there are 24! such reorderings of these items.

Putting this together, the probability of a random rearrangement having $ABC$ consecutively is

$$\frac{24!}{26!} = \frac{1}{26 \cdot 25} = \frac{1}{650}$$

Another way to do this problem is to use the multiplication rule. Think of generating a random reordering of the alphabet by randomly placing $A$, then $B$, then $C$, $D$, etc. The probability that $A$ gets placed somewhere where $BC$ could potentially follow is $24/26$ since anywhere but the last two positions could work. Then, assuming that happens, the chance of placing $B$ in the position after $A$ is $1/25$ since there are 25 open positions and only one of them is immediately after $A$. Similarly, assuming we have $B$ right after $A$, the chance of placing $C$ immediately after $B$ is $1/24$. This gives the same result:

$$\frac{24}{26} \cdot \frac{1}{25} \cdot \frac{1}{24} = \frac{1}{26 \cdot 25} = \frac{1}{650}$$