Recall that, for single-variable functions, we had the important notion of the area under the curve. Given a function $y = f(x)$, we started by dividing the interval $a \leq x \leq b$ into $n$ equal intervals $x_i \leq x \leq x_{i+1}$ of length $\Delta x = x_{i+1} - x_i$. We then approximated the areas over each interval by rectangles with width $\Delta x$ and height $f(x^*_i)$ where $x^*_i$ could be chosen arbitrarily from the interval $x_i \leq x^*_i \leq x_{i+1}$ (popular choices are the left endpoint, right endpoint, and mid-point).

This gives the following picture:

We also obtained the estimate

$$\text{Area} \approx \sum_{i=1}^{n} f(x^*_i) \Delta x.$$ 

We recognized that the estimate becomes better and better as we decrease the width of the intervals, which corresponds to increasing the number of intervals $n$. Taking the limit as $n \to \infty$, we arrived at

$$\text{Area} = \int_{a}^{b} f(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x^*_i) \Delta x.$$
This raises the question of what the analogue for the area under the curve is for functions of multiple variables. If we consider a function of the form $z = f(x, y)$, we can quickly see that the notion we need is that of volume under a surface. That is, we want to determine the volume bound by the surface $f(x, y)$ and the $(x, y)$-plane:

If we want to approximate the volume under the surface over a rectangular domain $\mathcal{R} = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$, then the single variable intuition generalizes completely. We divide the interval $a \leq x \leq b$ into $m$ intervals $x_i \leq x \leq x_{i+1}$ where $\Delta x = x_{i+1} - x_i$, and then divide $c \leq y \leq d$ into $n$ intervals $y_j \leq y \leq y_{j+1}$ where $\Delta y = y_{j+1} - y_j$. This gives $m \times n$ rectangular patches. Above each individual patch the volume is approximated by the volume of a three-dimensional box:

$$\text{height} \times \text{width} \times \text{depth} = f(x_i^*, y_j^*) \Delta x \Delta y$$

where the points $x_i \leq x_i^* \leq x_{i+1}$ and $y_j \leq y_j^* \leq y_{j+1}$ in each patch can be chosen arbitrarily. The volume over the entire domain $a \leq x \leq b$ and $c \leq y \leq d$ can therefore be approximated by

$$\text{Volume} \approx \sum_{i=1}^m \sum_{j=1}^n f(x_i^*, y_j^*) \Delta x \Delta y.$$  \hspace{1cm} (1)

As with one-dimensional Riemann sums, if we take smaller and smaller patches, corresponding to higher values for $m$ and $n$ (i.e. more overall
patches), we obtain a better and better estimate. The limit as \( m, n \to \infty \)
gives us

\[
\text{Volume} = \int \int_{\mathcal{R}} f(x, y) \, dx \, dy = \lim_{m \to \infty} \lim_{n \to \infty} \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i^*, y_j^*) \, \Delta x \Delta y.
\]

This is called the **double integral** of \( f(x, y) \) over \( \mathcal{R} \).

We will concern ourselves with methods for computing double integrals for specific functions \( f(x, y) \) in future lectures. For now, we will concern ourselves with familiarizing ourselves with the approximation formula (1) and the interpretation of volume under a surface.

---

**Example 1**

Estimate the volume under the surface \( f(x, y) = xy \) over the region

\[
\mathcal{R} = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 2, \ 0 \leq y \leq 2\}
\]

using \( m = n = 2 \). Compare the estimates obtained computing the height with (a) the lower left-most point of each region, (b) the upper right-most point, and (c) the midpoint. Compare the results with the actual value Volume = 4.

**Solution:** We can determine \( \Delta x \) and \( \Delta y \) by noticing that the length of the each interval (\( \Delta x \) and \( \Delta y \), respectively) can be given by the overall length (\( b - a \) and \( d - c \), respectively) divided by the number of subintervals (\( n \) and \( m \), respectively). This gives

\[
\Delta x = \frac{b - a}{m} = \frac{2 - 0}{2} = 1
\]

\[
\Delta y = \frac{d - c}{n} = \frac{2 - 0}{2} = 1.
\]

This corresponds to the following division of the domain:
The lower left-most points are labeled in green, the upper right-most points are labeled in red, and the mid-points are labeled in blue.

For the lower left-most points, we have

\[
\text{Volume} \approx \sum_{i=1}^{2} \sum_{j=1}^{2} f(x_i^*, y_j^*) \Delta x \Delta y \\
= [f(0, 0) + f(1, 0) + f(0, 1) + f(1, 1)] (1)(1) \\
= [(0)(0) + (1)(0) + (0)(1) + (1)(1)] (1)(1) \\
= (0 + 0 + 0 + 1)(1) = 1.
\]

For upper right-most points, we have

\[
\text{Volume} \approx \sum_{i=1}^{2} \sum_{j=1}^{2} f(x_i^*, y_j^*) \Delta x \Delta y \\
= [f(1, 1) + f(2, 1) + f(1, 2) + f(2, 2)] (1)(1) \\
= [(1)(1) + (2)(1) + (1)(2) + (2)(2)] (1)(1) \\
= (1 + 2 + 2 + 4)(1) = 9.
\]
For the mid-points, we have

\[
\text{Volume} \approx \sum_{i=1}^{2} \sum_{j=1}^{2} f(x_i^*, y_j^*) \Delta x \Delta y
\]

\[
= [f(0.5, 0.5) + f(1.5, 0.5) + f(0.5, 1.5) + f(1.5, 1.5)] (1)(1)
\]

\[
= [(0.5)(0.5) + (1.5)(0.5) + (0.5)(1.5) + (1.5)(1.5)] (1.5)(1.5)
\]

\[
= (0.25 + 0.75 + 0.75 + 2.25)(1) = 4.
\]

We can see that the lower left-most points significantly under-estimate the volume, while the upper right-upper points significantly over-estimate the volume. The mid-points happen to give the exact answer, although this is a coincidence. The pattern of over and under estimation should be expected given that \( f(x, y) = xy \) is monotonically increasing for \( x \geq 0 \) and \( y \geq 0 \):

Example 2

Estimate the volume bound by the \((x, y)\)-plane and the cylinder \( z^2 + y^2 = 1 \) over the domain

\[
\mathcal{R} = \{ (x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 2, \ -1 \leq y \leq 1 \}
\]

using \( m = 4 \) and \( n = 2 \) and the mid-point rule.

Solution: This corresponds to the volume below the half-cylinder

\[
z = f(x, y) = \sqrt{1 - y^2}.
\]
To determine how we carve up the domain, we notice that
\[ \Delta x = \frac{b - a}{m} = \frac{(2) - (0)}{4} = 0.5 \]
\[ \Delta y = \frac{d - c}{n} = \frac{(1) - (-1)}{2} = 1. \]

That is, to obtain 4 subintervals in \( x \) we need to choose subintervals of width \( \Delta x = 0.5 \), while obtaining 2 subintervals in \( y \) we need to choose subintervals of width \( \Delta y = 1 \). We can quickly determine that this gives \( m \times n = (4)(2) = 8 \) overall domain regions, each with area of width \( \times \) depth = \( \Delta x \Delta y = (0.5)(1) = 0.5 \).

To apply the midpoint rule, we divide the intervals in both direction in half to get the following set of 8 points:

\[ (0.25, -0.5), (0.75, -0.5), (1.25, -0.5), (1.75, -0.5), \]
\[ (0.25, 0.5), (0.75, 0.5), (1.25, 0.5), (1.75, 0.5). \]

The volume can then be approximated as

\[
\text{Volume} \approx \sum_{i=1}^{4} \sum_{j=1}^{2} f(x_i^*, y_j^*) \Delta x \Delta y
\]
\[
= (f(0.25, -0.5) + f(0.75, -0.5) + f(1.25, -0.5) + f(1.75, -0.5))
+ f(0.25, 0.5) + f(0.75, 0.5) + f(1.25, 0.5) + f(1.75, 0.5)) (0.5)(1)
\]
\[
= \left( 4\sqrt{1 - (0.5)^2} + 4\sqrt{1 - (-0.5)^2} \right) (0.5)
\]
\[
= 8\sqrt{\frac{3}{4}} (0.5) = 2\sqrt{3} \approx 3.464101616
\]

This compares with the true value of \( \pi \approx 3.141592654 \), which can be obtained by taking one half of the volume of a cylinder \( V = \pi r^2 h \) with radius \( r = 1 \) and height \( h = 2 \).
Example 3

Use a computer to redo Example 2 with the following values: (a) \(m = n = 5\); (b) \(m = n = 10\); (c) \(m = n = 100\); and (d) \(m = n = 500\).

Solution: Notice that it would be unwieldy to do these computations by hand. Even the smallest case has \(m \times n = (5)(5) = 25\) regions. That is a lot of computation!

Our computer, however, does not care. We can quickly have it compute that:

<table>
<thead>
<tr>
<th>(m)</th>
<th>(n)</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>3.464101616</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>3.22642422</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>3.171987824</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>3.142565555</td>
</tr>
<tr>
<td>500</td>
<td>500</td>
<td>3.141679764</td>
</tr>
</tbody>
</table>

We can see that—as expected!—the estimate converges to the true value \(\text{Volume} = \pi\) as we take finer and finer divisions of the region \(\mathcal{R}\).
Section 2: Averages

An immediate application of double integrals is computing averages. For instance, we may know the temperature distribution over a region and wish to compute the average temperature in the region. This data may come in the form of sample points or as a function \( T(x, y) \).

To accomplish the process of averaging, we need to compute the value per unit of area. This gives the approximate

\[
\text{Average} \approx \frac{1}{A(R)} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_i^*, y_j^*) \Delta x \Delta y
\]

where \( A(R) \) is the area of \( R \). For rectangular regions, we have \( A(R) = (b-a)(d-c) \). Taking the limit at \( n \to \infty \), we have

\[
\text{Average} = \frac{1}{A(R)} \iint_{R} f(x, y) \, dx \, dy.
\]

We will not concern ourselves with exactly computing double integrals yet.

Example 4

Estimate the average height of the function \( f(x, y) = x^2y^2 - x^2 - y^2 + 1 \) over the region

\[
R = \{ (x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1, \, -1 \leq y \leq 1 \}
\]

with \( n = m = 2 \) and the midpoint rule.

Solution: The only difference between this and the previous questions is that we need to correct by the area of the region \( A(R) \). We clearly have \( \Delta x = 1 \) and \( \Delta y = 1 \) so that the midpoint rule gives

\[
\text{Volume} \approx (f(-0.5, -0.5) + f(-0.5, 0.5) + f(0.5, -0.5) + f(0.5, 0.5)) \Delta x \Delta y
\]

\[
= 4 \left( \left( \frac{1}{4} \right) \left( \frac{1}{4} \right) - \left( \frac{1}{4} \right) - \left( \frac{1}{4} \right) + 1 \right) (1)(1) = \frac{9}{4}.
\]
We have that the area of the region is given by

\[ A(\mathcal{R}) = (b - a)(d - c) = (2)(2) = 4 \]

so that the average value of the function over \( \mathcal{R} \) is given by

\[
\text{Average} = \frac{1}{A(\mathcal{R})} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x^*_i, y^*_j) \Delta x \Delta y = \frac{1}{4} \cdot \frac{9}{4} = \frac{9}{16}.
\]

The result can be interpreted as taking the function (which looks like a hill) and imagining chopping off the summit and using it to fill in the valleys beneath it, until everything is completely level. The volume (i.e. amount of stuff) is conserved, but the intricacies of the terrain is lost. See below:

Example 5

Suppose that the temperature in a region \( \mathcal{R} = [0, \pi] \times [0, \pi] \) is given in Fahrenheit by

\[ T(x, y) = 5 \sin(x) + 5 \cos(x) + 70. \]

Use four equally sized rectangles and the mid-point rule to estimate the average temperature in the region.

**Solution:** We can quickly determine that \( \Delta x = \Delta y = \pi/2 \) so that
the midpoints are \((\pi/4, \pi/4), (3\pi/4, \pi/4), (\pi/4, 3\pi/4)\) and \((3\pi/4, 3\pi/4)\).

The area of the region is \(A(\mathcal{R}) = \pi^2\). It follows that

\[
\text{Average} = \frac{1}{A(\mathcal{R})} \sum_{i=1}^{2} \sum_{j=1}^{2} f(x^*_i, y^*_j) \Delta x \Delta y
\]

\[
= \frac{1}{\pi^2} \left[ \left( 5 \sin\left(\frac{\pi}{4}\right) + 5 \cos\left(\frac{\pi}{4}\right) + 70 \right) + \left( 5 \sin\left(\frac{3\pi}{4}\right) + 5 \cos\left(\frac{\pi}{4}\right) + 70 \right) + \left( 5 \sin\left(\frac{\pi}{4}\right) + 5 \cos\left(\frac{3\pi}{4}\right) + 70 \right) + \left( 5 \sin\left(\frac{3\pi}{4}\right) + 5 \cos\left(\frac{3\pi}{4}\right) + 70 \right) \right] \left(\frac{\pi}{2}\right)^2
\]

\[
= \frac{5}{\sqrt{2}} + 70 \approx 73.54.
\]

So the average temperature in the region is approximately 73.54°F.