10 Integration by parts; numerical integration

10.1 Integration by parts

Integration by parts is just the **product rule applied to integration**. To introduce this we start with the product rule and integrate:

Product rule: \[(uv)' = uv' + u'v\]

Integrate both sides: \[\int (uv)' \, dx = \int uv' \, dx + \int u'v \, dx\]

Since \(\int (uv)' = uv\):
\[uv = \int uv' \, dx + \int u'v \, dx\]

Simple algebra:
\[\int u'v \, dx = uv - \int uv' \, dx\]

We summarize by putting the last equation in a box: the **integration by parts formula** is

\[
\int uv' \, dx = uv - \int u'v \, dx
\]

**Example 10.1.** Find \(\int xe^x \, dx\).

**answer:** (I suggest you learn to use the following format.)

Make the following choices for \(u\) and \(v'\):

\[
\begin{array}{c|c}
\text{u} & \text{v'} \\
\hline
x & e^x \\
\end{array}
\]

The integration by parts formula says
\[
\int xe^x \, dx = xe^x - \int e^x \, dx = xe^x - e^x + C.
\]

**Discussion:** Choosing \(u\) and \(v'\) is an art. Notice that in the above example \(u = x\) gets simpler when we take \(u'\) and that \(v' = e^x\) is easy to integrate.

**Example 10.2.** Find \(\int \ln(x) \, dx\).

**answer:** At first glance this integral does not look like a candidate for integration by parts because it seems to have only one part. But we can always let \(v' = 1\) for the second part:

\[
\begin{array}{c|c}
\text{u} & \text{v'} \\
\hline
\ln x & 1 \\
\frac{1}{x} & x \\
\end{array}
\]

So,
\[
\int \ln x \, dx = x \ln x - \int dx = x \ln x - x + C.
\]
Example 10.3. (Sometimes you have to iterate.) Find $\int x^2 e^x \, dx$.

**Answer:** Choose $u$ and $v'$:

\[
\begin{align*}
  u &= x^2 \quad v' = e^x \\
  u' &= 2x \quad v = e^x
\end{align*}
\]

Thus,

\[
\int x^2 e^x \, dx = x^2 e^x - \int 2xe^x \, dx.
\]

Now we’ve reduced the original integral to another integral (with a smaller power of $x$) that we have to integrate by parts. Actually, we did this new integral in a previous example:

\[
\int 2xe^x \, dx = 2xe^x - 2e^x.
\]

So, the answer to the question is $\int x^2 e^x \, dx = x^2 e^x - 2xe^x + 2e^x + C$.

Example 10.4. (Another trick for your bag of tricks.) Find $\int e^x \cos(x) \, dx$.

**Answer:** Both parts $e^x$ and $\cos(x)$ are easy to integrate and differentiate so we make a guess at how to split them:

\[
\begin{align*}
  u &= \cos(x) \quad v' = e^x \\
  u' &= -\sin(x) \quad v = e^x
\end{align*}
\]

This gives

\[
\int e^x \cos(x) \, dx = e^x \cos(x) + \int e^x \sin(x) \, dx.
\]

Here’s the trick: the integral on the right is of the same form as before so we try a second integration by parts:

\[
\begin{align*}
  u &= \sin(x) \quad v' = e^x \\
  u' &= \cos(x) \quad v = e^x
\end{align*}
\]

Which gives

\[
\int e^x \cos(x) \, dx = e^x \cos(x) + \int e^x \sin(x) \, dx = e^x \cos(x) + e^x \sin(x) - \int e^x \cos(x) \, dx.
\]

Now the integral on the right is the same as our original integral with a minus sign. Algebra now yields:

\[
2 \int e^x \cos(x) \, dx = e^x \cos(x) + e^x \sin(x).
\]

So the answer to the question is:

\[
\int e^x \cos(x) \, dx = \frac{e^x(\cos(x) + \sin(x))}{2}.
\]

Example 10.5. (Definite integrals.) Compute $\int_0^\pi x \sin x \, dx$.

**Answer:** Using integration by parts

\[
\begin{align*}
  u &= x \quad v' = \sin(x) \\
  u' &= 1 \quad v = -\cos(x)
\end{align*}
\]
we get
\[ \int_0^\pi x \sin(x) \, dx = -x \cos(x)|_0^\pi + \int_0^\pi \cos(x) \, dx = \pi + \sin(x)|_0^\pi = \pi. \]
Note well that both terms from the integration by parts get limits of integration. Remember that integration by parts helps you find the antiderivative and that for a definite integral you have to put the limits on the entire antiderivative.

**Inverse trig functions**

We know that \( \frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}}. \) (We’ll remind you how to find this in a minute.)

Using this formula we can use integration by parts to find \( \int \sin^{-1}(x) \, dx. \)

Choose

\[
\begin{array}{c|c}
 u &= \sin^{-1} x & v' &= 1 \\
 u' &= \frac{1}{\sqrt{1-x^2}} & v &= x \\
\end{array}
\]

This gives

\[
\int \sin^{-1}(x) \, dx = x \sin^{-1}(x) - \int \frac{x}{\sqrt{1-x^2}} \, dx = x \sin^{-1}(x) + \sqrt{1-x^2} + C.
\]

Let’s now recall how to compute the derivative of \( y = \sin^{-1}(x); \)

We have \( \sin(y) = x. \) We can take the derivative of both sides with respect to \( x \) using the chain rule (remember \( y \) is a function of \( x \)) to get \( \cos(y) \frac{dy}{dx} = 1. \) Solving for \( \frac{dy}{dx} \) we get

\[
\frac{dy}{dx} = \frac{1}{\cos(y)} = \frac{1}{\sqrt{1-x^2}}.
\]

The value \( \cos(y) = \sqrt{1-x^2} \) was found using the trig-triangle above.

**A reduction formula.** Reduction formulas are general formulas that show how to reduce an integral with a power to a similar one with a smaller power. Here we will show how to find and use the following reduction formula.

\[
\int \sin^n(x) \, dx = \frac{1}{n} \sin^{n-1}(x) \cos(x) + \frac{n-1}{n} \int \sin^{n-2}(x) \, dx. \tag{1}
\]

Notice that the left-hand side involves an integral with power \( n \) and the right-hand side has an integral with the smaller power \( n - 2. \)
Before proving the reduction formula \ref{1} we show how to use it in an example.

**Example 10.6.** Find \( \int \sin^5(x) \, dx \).

**Answer:** We use the reduction formula \ref{1} repeatedly.

\[
\int \sin^5(x) \, dx = -\frac{1}{5} \sin^4(x) \cos(x) + \frac{4}{5} \int \sin^3(x) \, dx \\
= -\frac{1}{5} \sin^4(x) \cos(x) - \frac{4}{5} \cdot \frac{1}{3} \sin^2(x) \cos(x) + \frac{4}{5} \cdot \frac{2}{3} \int \sin(x) \, dx \\
= -\frac{1}{5} \sin^4(x) \cos(x) - \frac{4}{5} \cdot \frac{1}{3} \sin^2(x) \cos(x) - \frac{4}{5} \cdot \frac{2}{3} \cos(x) + C.
\]

**Proof of \ref{1}** Using integration by parts choose \( u \) and \( v' \) by

\[
\begin{align*}
    u &= \sin^{n-1}(x) \\
    u' &= (n-1) \sin^{n-2}(x) \cos(x) \\
    v &= \sin(x) \\
    v' &= \cos(x)
\end{align*}
\]

This gives

\[
\int \sin^n(x) \, dx = -\sin^{n-1}(x) \cos(x) + (n-1) \int \sin^{n-2}(x) \cos^2(x) \, dx \\
= -\sin^{n-1}(x) \cos(x) + (n-1) \int \sin^{n-2}(x) (1 - \sin^2(x)) \, dx \\
= -\sin^{n-1}(x) \cos(x) + (n-1) \int \sin^{n-2}(x) - \sin^n(x) \, dx.
\]

Moving the \( \sin^n(x) \) term on the right side to the left side we get

\[
n \int \sin^n(x) \, dx = -\sin^{n-1}(x) \cos(x) + (n-1) \int \sin^{n-2}(x) \, dx.
\]

Now dividing by \( n \) gives the reduction formula.

### 10.2 Numerical Integration

**Numerical integration** refers to a number of different ways to approximate the value of a definite integral. There are several reasons we might want to use numerical integration:

- We can use computers to evaluate the integrals. With modern computers we can easily do a calculation with tens of thousands or even millions of terms. This allows us to make extremely accurate approximations of integrals.

- For many functions there is no closed formula for the antiderivative so we are forced to turn to numerical integration to find the value. For example,

\[
\int_{-\infty}^{\infty} e^{-x^2} \, dx, \quad \int_{0}^{\pi} \sqrt{\sin(x)} \, dx, \quad \text{and} \quad \int_{0}^{2\pi} \sqrt{1 + \sin^2(x)} \, dx
\]

are all integrals which can’t be computed by finding an antiderivative, so we must do some type of numerical approximation.
Riemann sums. When we first defined definite integrals we introduced the notion of Riemann sums. To make a Riemann sum you divide the interval of integration into $n$ subintervals and sum the areas of rectangles above each of these subintervals. There are many possible ways to choose the rectangles above each interval:

- **Right Riemann sum (RRS):**

  $$\int_{a}^{b} f(x) \, dx = \text{area} \approx (y_1 + y_2 + \ldots + y_n) \Delta x = \sum_{j=1}^{n} y_j \Delta x.$$  

  ![Figure showing the right Riemann sum approximating an integral.]

  For the right Riemann sum the height of a rectangle above an interval is the height of the curve above the right endpoint of the interval. Notice the right Riemann sum starts at $j = 1$ and ends at $j = n$.

- **Left Riemann sum (LRS):**

  $$\int_{a}^{b} f(x) \, dx = \text{area} \approx (y_0 + y_1 + \ldots + y_{n-1}) \Delta x = \sum_{j=0}^{n-1} y_j \Delta x.$$  

  ![Figure showing the left Riemann sum approximating an integral.]

  For the left Riemann sum the height of a rectangle above an interval is the height of the curve above the left endpoint of that interval. Notice that the left Riemann sum starts at $j = 0$ and ends at $j = n - 1$.

- **Mid Riemann sum:** For the mid Riemann sum the height of a rectangle above an interval is the height of the curve above the midpoint of that interval.
• **Maximum Riemann sum:** For the maximum Riemann sum the height of a rectangle above an interval is the height of the highest point on the curve above that interval.

• **Minimum Riemann sum:** For the minimum Riemann sum the height of a rectangle is the height of the lowest point on the curve above that interval.

• **Random Riemann sum:** For the random Riemann sum the height of a rectangle above an interval is the height of a random point on the curve above that interval.

**Example 10.7.** Estimate \( \int_0^1 \sqrt{1-x^3} \, dx \) using right and left Riemann sums.

**Answer:** To avoid too much calculation we’ll do this with \( n = 4 \). On a computer we could use a much bigger value of \( n \). Here is one good way to organize the calculation in a table.

<table>
<thead>
<tr>
<th>( j )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_j )</td>
<td>0</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>1</td>
</tr>
<tr>
<td>( y_j )</td>
<td>1</td>
<td>0.992</td>
<td>0.935</td>
<td>0.760</td>
<td>0</td>
</tr>
</tbody>
</table>

We first compute \( \Delta x = \frac{b-a}{n} = \frac{1}{4} \). Now the two Riemann sums are

- Right Riemann sum (RRS) = \( (y_1 + y_2 + y_3 + y_4) \Delta x \approx 0.672 \)
- Left Riemann sum (LRS) = \( (y_0 + y_1 + y_2 + y_3) \Delta x \approx 0.922 \).

Notice that for \( n = 4 \) our two approximations are fairly different. Using a computer we can compute these sums for much larger \( n \):

- With \( n = 100 \): RRS = 0.836, LRS = 0.846.
- With \( n = 500 \): RRS = 0.840, LRS = 0.842.

We can see the values of the estimates converging as \( n \) gets bigger. The “true” value of the integral = 0.8413. (This was computed numerically using Simpson’s rule with \( n = 1.6 \times 10^6 \). We will learn about Simpson’s rule shortly.)

**Trapezoidal Rule:** A more accurate rule than the ones using rectangles is the trapezoidal rule. The idea here is to use trapezoids above each interval in place of rectangles. The top of the trapezoid is found by connecting the points the two endpoints on the part of the curve above the interval –see the figure below. Notice how well the trapezoids approximate the area.

![Figure showing the trapezoidal rule for approximating an integral.](image)

Figure showing the **trapezoidal rule** for approximating an integral.

We know that the area of a trapezoid is

\[
\text{base} \times \text{average of the legs}
\]
Summing the area of all $n$ trapezoids we get

$$\int_a^b f(x) \, dx \approx \left( \frac{y_0 + y_1}{2} + \frac{y_1 + y_2}{2} + \cdots + \frac{y_{n-1} + y_n}{2} \right) \Delta x = \left( \frac{1}{2} y_0 + y_1 + y_2 + \cdots + y_{n-1} + \frac{1}{2} y_n \right) \Delta x.$$

Another way to describe the trapezoidal rule is that it’s the average of left and right Riemann sums.

**Example 10.8.** (Same integral as above.) Approximate the value of $\int_0^1 \sqrt{1-x^3} \, dx$ using the trapezoidal rule with $n = 4$.

**answer:** We can use the table giving $y_0, \ldots, y_4$ from the previous example.

$$\int_0^1 \sqrt{1-x^3} \, dx \approx \left( \frac{1}{2} \cdot 1 + 0.992 + 0.935 + 0.760 + \frac{1}{2} \cdot 0 \right) \cdot 0.25 = 0.797.$$

With $n = 30$, the trapezoidal rule approximates this integral as 0.839. The trapezoidal rule converges to the true value much faster, i.e. with smaller $n$, than the rectangle rules.

**Simpson’s Rule:** We can get an even more accurate approximation by using parabolic caps on the area above each interval. This gives us Simpson’s rule. You can read the derivation in the textbook.

The rule requires that $n$ be even. The formula is

$$\int_a^b f(x) \, dx \approx \frac{1}{3} \left( y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 4y_{n-1} + y_n \right) \Delta x.$$

(Note: there are other ways to state Simpson’s rule that don’t require $n$ to be even. We will be consistent in using the above formulation.

**Example 10.9.** (Same integral as above.) Approximate $\int_0^1 \sqrt{1-x^3} \, dx$ using Simpson’s rule with $n = 4$.

**answer:** We can once again use the table of values in the previous examples:

$$\int_0^1 \sqrt{1-x^3} \, dx \approx \frac{1}{3} \left( 1 + 4(0.992) + 2(0.935) + 4(0.760) + 0 \right) \cdot 0.25 = 0.823.$$

Notice how even with $n = 4$ Simpson’s rule gives a pretty good approximation of the true value of 0.843.

With $n = 30$: Simpson’s rule gives $\int_0^1 \sqrt{1-x^3} \approx 0.840$.

**Approximating sums using integrals.** As we’ve done before we can turn approximation on its head and approximate a sum using an integral.

**Example 10.10.** Use the trapezoidal rule to estimate $1^2 + 2^2 + 3^2 + \cdots + 100^2$. Is the estimate too high or too low?

**answer:** The trapezoidal rules says

$$\int_0^{100} x^2 \, dx \approx \frac{0^2}{2} + 1^2 + 2^2 + \cdots + 99^2 + \frac{100^2}{2}.$$
We prefer to use the trapezoidal rule to approximate the integral because it is more accurate. The price we pay is that the above sum is not exactly the sum we want to estimate. This is okay, often in problem solving you can get something similar to what you want and then tweak it to get exactly what you want. In our case we can tweak the trapezoidal approximation by adding $\frac{100^2}{2}$ to both sides:

$$\int_0^{100} x^2 \, dx + \frac{100^2}{2} \approx \frac{0^2}{2} + 1^2 + 2^2 + \ldots + 99^2 + \frac{100^2}{2} + \frac{100^2}{2} = 1^2 + \ldots + 100^2.$$  

The expression on the right is exactly the sum we want to approximate. The expression left is easy to compute

$$\int_0^{100} x^2 \, dx + \frac{100^2}{2} = \frac{10^6}{3} + \frac{10^4}{2} = 338333.33.$$  

So our approximation is $1^2 + 2^2 + \ldots + 100^2 \approx 338333.33$. (The exact sum is 338350).

Since $x^2$ is concave up each of the trapezoids is above the curve. This means the sum is bigger than the integral used to make the estimate, so the estimate is too low.