Project 3 - Numerical integration

Mari Dahl Eggen

October 18, 2015

Abstract

In this project we will look at and try out four different methods of numerical integration. The methods is named Gaussian Legendre, Gaussian Laguerre, brute force Monte Carlo, and Monte Carlo with importance sampling. The methods will be compared with each other, and we will see that there is a big difference in accuracy and effectiveness between the methods.

Introduction

The integral one have to solve to find the quantum mechanical expectation value of the correlation energy between two electrons which repel each other via the classical Coulomb interaction, has a closed form solution. In the following we will solve this integral numerically with four different integration methods, and compare the results we get with the analytically calculated value. From that we can consider how good each of the integration methods is, and we can compare them with each other both in how accurate results they give, and what the execution time for the method is. Some times a method can give a quite good accuracy, but the time it takes to get that value is longer than we want wait for the answer.

1 Theory

1.1 Quantum mechanical expectation value

We define $\mathbf{r}_i = x_i \mathbf{e}_x + y_i \mathbf{e}_y + z_i \mathbf{e}_z$, and though $r_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$. Then the ansatz for the wave function for two electrons is given by

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = e^{-\alpha(r_1+r_2)}.$$

In this case we want to solve the quantum mechanical expectation value of the correlation energy between two electrons which repel each other via the classical Coulomb interaction, that is

$$\langle \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \rangle = \int_{-\infty}^{\infty} d\mathbf{r}_1 d\mathbf{r}_2 e^{-2\alpha(r_1+r_2)} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{1}$$

We will in this case fix $\alpha = 2$, which corresponds to the charge of the helium atom, namely $Z = 2$. Then this integral can be solved in closed form, and the answer is $\frac{5\pi}{16^2}$. We rewrite Eq. (1) to spherical coordinates, and get

$$d\mathbf{r}_1 d\mathbf{r}_2 = r_1^2 dr_1 r_2^2 dr_2 \sin\theta_1 d\theta_1 \sin\theta_2 d\theta_2 d\phi_1 d\phi_2,$$

$$\frac{1}{r_{12}} = \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 cos(\beta)}}$$

and

$$cos(\beta) = cos(\theta_1)cos(\theta_2) + sin(\theta_1)sin(\theta_2)cos(\phi_1 - \phi_2)).$$

1.2 Legendre

The Legendre method has the weight function $W(x) = 1$ on the interval $x \in [-1, 1]$. We have that

$$I = \int_{-1}^{1} f(x)dx = \int_{-1}^{1} W(x)g(x)dx = \int_{-1}^{1} g(x)dx = \sum_{i=1}^{N} \omega_i g(x_i),$$

where the weights and points can be found by the Legendre method.

1.3 Laguerre

The Laguerre method has the weight function $W(x) = x^\alpha e^{-x}$ on the interval $0 \leq x \leq \infty$. We have that

$$I = \int_0^\infty f(x)dx = \int_0^\infty W(x)g(x)dx = \int_0^\infty x^\alpha e^{-x}g(x)dx = \sum_{i=1}^N \omega_i g(x_i).$$

where the weights and points can be found by the Laguerre method. To obtain the form on the integral in Eq. (1), so that we can use the Laguerre method to solve it, we have to do a substitution on the spherical form.

$$\int_0^\infty \int_0^\infty \int_0^\pi \int_0^\pi \int_0^{2\pi} \int_0^{2\pi} r_1^2 r_2^2 e^{-4r_1} e^{-4r_2} \frac{\sin\theta_1 \sin\theta_2}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\beta}} dr_1 dr_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2$$

$$(2)$$

We substitute $u_1 = 4r_1$ and $u_2 = 4r_2$, and get

$$\int_0^\infty \int_0^\infty \int_0^\pi \int_0^\pi \int_0^{2\pi} \int_0^{2\pi} \frac{u_1^2}{16}\frac{u_2^2}{16} e^{-u_1} e^{-u_2} \frac{4\sin\theta_1 \sin\theta_2}{\sqrt{u_1^2 + u_2^2 - 2u_1 u_2 \cos\beta}} \frac{du_1}{4} \frac{du_2}{4} d\theta_1 d\theta_2 d\phi_1 d\phi_2$$

$$= \frac{1}{1024} \int_0^\infty \int_0^\infty \int_0^\pi \int_0^\pi \int_0^{2\pi} \int_0^{2\pi} u_1^2 u_2^2 e^{-u_1} e^{-u_2} \frac{\sin\theta_1 \sin\theta_2}{\sqrt{u_1^2 + u_2^2 - 2u_1 r_2 \cos\beta}} du_1 du_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2$$

$$(3)$$

$$= \frac{1}{1024} \int_0^\infty \int_0^\infty \int_0^\pi \int_0^\pi \int_0^{2\pi} \int_0^{2\pi} W(u_1)W(u_2)g(u,\theta,\phi)du_1 du_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2.$$

1.4 Monte Carlo - brute force

With brute force Monte Carlo we do simulations with random numbers generated from an uniform distribution. For a uniform distribution we have

$$p(x)dx = \begin{cases} dx & \text{if } 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases},$$

where $p(x) = 1$, and where $\int_{-\infty}^{\infty} p(x)dx = 1$ is satisfied. Then we have that

$$I = \int_0^1 p(x)f(x) \simeq \sum_{i=1}^N f(x) = \langle f \rangle,$$

where N is the number of simulations. If the boundaries of the integral we want to evaluate is different from the limits in the uniform distribution, we also have to do a change of variables. We find the formula for change of variables from a uniform distribution with general limits.

$$p(y)dy = \begin{cases} \frac{dy}{b-a} & \text{if } a \leq y \leq b \\ 0 & \text{else} \end{cases}$$

Because of conservation of probability we have

$$p(y)dy = \frac{dy}{b-a} = dx,$$

By integration we find x expressed by y, and by inverting we find the formula for change of variables.

$$x(y) = \int_a^y \frac{dy'}{b-a} = \frac{y}{b-a} - \frac{a}{b-a}$$

$$\Rightarrow \quad y(x) = a + (b-a)x$$

1.5 Monte Carlo - importance sampling with exponential distribution

In general important sampling means that we find a PDF $p(y)$ whose behavior resembles that of the function $f(x)$ that we want to find the integral of, in a given interval $[a, b]$. We need to make sure that the normalization condition $\int_a^b p(y)dy = 1$ is fulfilled. We can then rewrite the integral as

$$I = \int_a^b f(y)dy = \int_a^b p(y)\frac{f(y)}{p(y)}dy.$$

Since the random numbers are generated from a uniform distribution $p(x)$ with $x \in [0, 1]$, we need to perform a change of variables from x to y. This is done in the same way as discussed in the previous section. Then we finally get

$$I = \int_{\tilde{a}}^{\tilde{b}} \frac{f(y(x))}{p(y(x))} dx \simeq \frac{1}{N} \sum_{i=1}^{N} \frac{f(y(x_i))}{p(y(x_i))}.$$

In our case we can use a PDF on the form of an exponential distribution $p(y) = e^{-y}$. If we look at the expression in Eq. (3), we see that the factors e^{-u_1} and e^{-u_2} corresponds to the exponential distribution. We then choose to use the PDF

$$p(r) = Ae^{-4r},$$

where r corresponds to both r_1 and r_2, and A is a normalization constant. A is found by the normalization condition.

$$A \int_0^\infty e^{-4r} dr = A \left[\frac{-e^{-4r}}{4} \right]_0^\infty = -\frac{A}{4}(0-1) = \frac{A}{4} = 1 \quad \Rightarrow \quad A = 4$$

We then have that $p(r) = 4e^{-4r}$, and we now want to find the formula for the change of variables from x, generated by the uniform distribution, to r.

$$dx = p(r)dr \quad \Rightarrow \quad x(r) = 4 \int_0^r e^{-4r'} dr' = 4 \left[\frac{-e^{-4r'}}{4} \right]_0^r$$

$$= -(e^{-4r} - e^0) = 1 - e^{-4r} \quad \Rightarrow \quad r(x) = \frac{-\ln(1-x)}{4},$$

where $x \in [0, 1]$ and $r \in [0, \infty)$.

2 Method

2.1 Implementation of integrand

When we implement the algorithm for solving the integral, we need to account for the problems that occur when the denominator $|\mathbf{r}_1 - \mathbf{r}_2| = 0$. In the denominator there is also taken a square root, that will cause problems when the expression inside the square root is less that zero. We choose to discard all terms where the denominator is close to zero and all the terms that turns out to be imaginary, and implements the test in Listing 1 that is provides this.

Listing 1: If-test that checks if the denominator in the integrand evaluated is close to zero or if the integrand turns out to be imaginary.

```
if(deno < pow(10.,-6.) || root < 0.0) { return 0;}
```

2.2 Legendre

In order to use the Legendre method to evaluate the integral in Eq. (1), we have to reduce the limits. We need to find out when the wave function goes to zero.

$$e^{-4r_i} = 0.00001 \quad \Rightarrow \quad -4r_i = \ln(0.00001) \quad \Rightarrow \quad r_i \simeq \frac{11.51}{4} \simeq 2.89$$

If we set the limits to $r \in [-3.0, 3.0]$, the error due to reduced limits is negligible. This is also confirmed by the plot of the wave function for two electrons in Figure 1. We can now use the Legendre method for the integral in Eq. (1).

2.3 Laguerre

The Laguerre polynomial is defined for $x \in [0, \inf)$. We can therefore evaluate our integral with this method if we use the integral on its spherical form, as shown in Eq. (2).

2.4 Monte Carlo

When we evaluate the integral with brute force Monte Carlo, we can use the integral as it is implemented in Eq. (1), because the PDF $p(x) = 1$. For the importance sampling method we need to use the exponential distribution to get the best results, and therefore we use the integral as it is shown in Eq. (3).
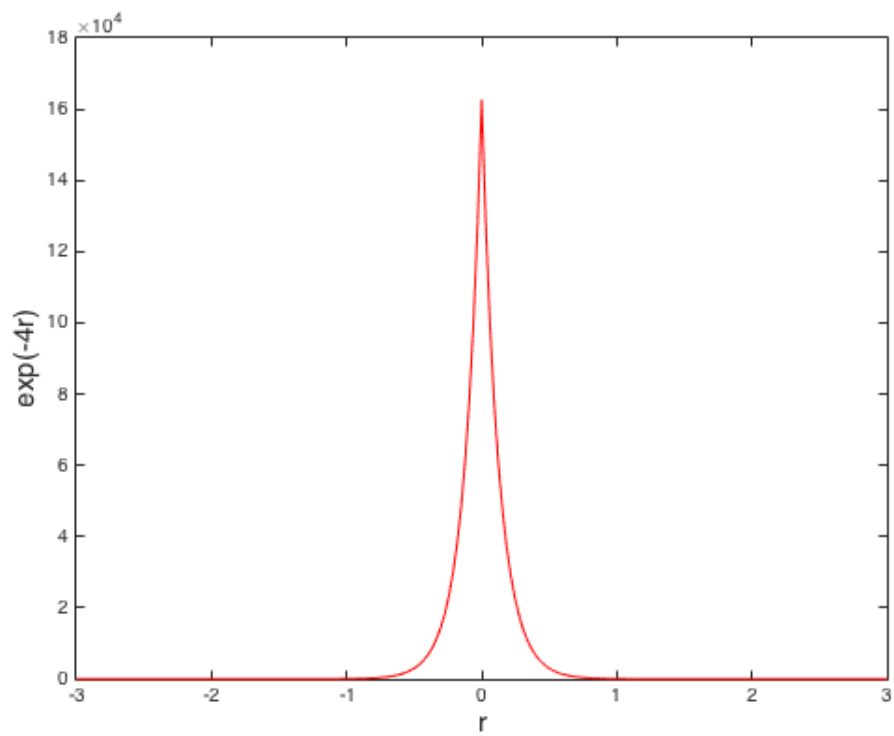
Figure 1: The wave function of two electrons.

3 Results and discussion

3.1 Legendre

From Table 1 we can see that the Legendre method is unstable and inaccurate for small values of N. When we set the numbers of integration steps to $N = 30$, the result is not so bad, but in exchange we have to wait for the result in almost 3 minutes.

N	Legendre	Relative error	Time spent [sec]
5	0.264249	0.37083	0.00453
10	0.0719797	0.626595	0.218
20	0.156139	0.190004	14.1
30	0.177283	0.080319	162

Table 1: The integral in Eq. (1) evaluated with the Legendre method for different values of integration steps N. The relative error of the results and the execution time is also included.

3.2 Laguerre

From Table 2 we can see that compared with Legendre, this method is much better. We can see that we reach about the same accuracy with Laguerre at $N = 10$, as we do with Legendre at $N = 30$. The big difference then, is that we only have to wait for about 0.2 seconds when we use Laguerre. We can see that the results can be even better by use of Languerre for bigger N, but the execution time is growing fast when N is increased.

N	Laguerre	Relative error	Time spent [sec]
5	0.17345	0.100205	0.00333
10	0.186457	0.0327256	0.191
20	0.191082	0.00873564	12.0
30	0.192114	0.00338234	143

Table 2: The integral in Eq. (2) evaluated with the Laguerre method for different values of integration steps N. The relative error of the results and the execution time is also included.

3.3 Monte Carlo - brute force

The analytically calculated result for the integral is $\frac{5\pi^2}{256} \simeq 0.192766$, and from Table 3 we can see that the results are getting close to the real value from $10e5$ simulations and up. We can see that the standard deviation values are incorrect, because they show the same magnitude in expected error for all numbers of N, which is completely wrong. I do not know why or where that error occurs. Notice that the numerically calculated value for $N = 10e5$ is better than that for $N = 10e6$. The reason for this is that the random number generator is fed by a random seed for every execution, so that the results differs in accuracy from one execution to another. In general the result is supposed to be more accurate for bigger number om simulations. The result for $10e7$ is close to the real value, but again we have to wait for over a minute to get that value.

3.4 Monte Carlo - importance sampling with exponential distribution

The results from this method is listed in Table 4. Here the standard deviation values complies with the accuracy in the numerically calculated values. We can see that the accuracy is extraordinary good already for $N = 10e4$, and it just keep getting better for a bigger number of simulations. This method has approximately the same execution time as the brute force method, but is giving remarkably better results for lower N. This means that we can get the same accuracy in the results from this method as the results from the brute force Monte Carlo method, in a much shorter time.

N	Brute force MC	Standard deviation	Time spent [sec]
$10e3$	0.094295605	0.078311202	0.00537
$10e4$	0.072395409	0.035428948	0.0409
$10e5$	0.20270969	0.057600078	0.392
$10e6$	0.25976629	0.067301031	4.32
$10e7$	0.19705999	0.012963037	40.0

Table 3: The integral in Eq. (1) evaluated with brute force Monte Carlo method, for different number of simulations N. The standard deviation of the results and the execution time is also included.

N	Importance sampling MC	Standard deviation	Time spent [sec]
$10e3$	0.22191239	0.029888926	0.00443
$10e4$	0.19899945	0.0055208307	0.0555
$10e5$	0.19296049	0.0010245359	0.408
$10e6$	0.19297351	0.00032495340	4.21
$10e7$	0.19271743	0.00010396595	41.1

Table 4: The integral in Eq. (2) evaluated with importance sampling Monte Carlo method, for different number of simulations N. The standard deviation of the results and the execution time is also included.

4 Conclusion