Models, Algorithms and Data (MAD): Introduction to Computing

Spring semester 2019

Prof. Dr. Jens Honore Walther
Dr. Julija Zavadlav
ETH Zentrum, CLT
CH-8092 Zürich

Set 11

Issued: 11.05.2019; Due: 19.05.2019

In this exercise, we will learn about (1) integrating in higher dimensions using Cartesian-products of 1D integrations and (2) Monte-Carlo integration. We will investigate the advantage of Monte-Carlo integration in high dimensions.

Question 1: Volume under surface

The surface shown in Fig. 1 is described by the following equation:

$$z(x, y) = \sin(x) + \cos(y) + 1 \tag{1}$$

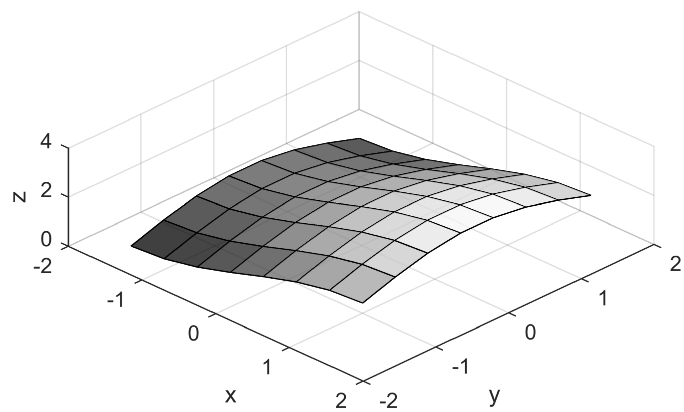

Figure 1: Surface described by Eq. 1.

a) Evaluate the volume under the surface by computing the following integral analytically:

$$V = \int_{y=-\pi/2}^{\pi/2} \int_{x=-\pi/2}^{\pi/2} z(x, y)\ dx\ dy$$

b) Write a program to compute V numerically. Use a 'Cartesian-product' of two one-dimensional trapezoidal rules.

   The Cartesian-product approach breaks up the 2D integral into sums of 1D integrals as follows:

$$V = \int_{y=-\pi/2}^{\pi/2} \int_{x=-\pi/2}^{\pi/2} z(x, y)\ dx\ dy = \sum_j \sum_i w_j\ w_i\ z(x_i, y_j), \tag{2}$$

where w_i and w_j are node-weights of the trapezoidal rule.

Fig. 2 provides an intuitive illustration of the Cartesian-product approach. We first consider a single 'slice' of the curve at $y_1 = -\pi/2$, and evaluate the area under this slice by dividing up the x-axis into sub-intervals. This operation can be represented by the following equation:

$$A_1 = \sum_i w_i z(x_i, y = -\pi/2)$$

Similarly, we compute the area of the slice at $y_2 = -\pi/2 + \pi/5$:

$$A_2 = \sum_i w_i z(x_i, y = -\pi/2 + \pi/5)$$

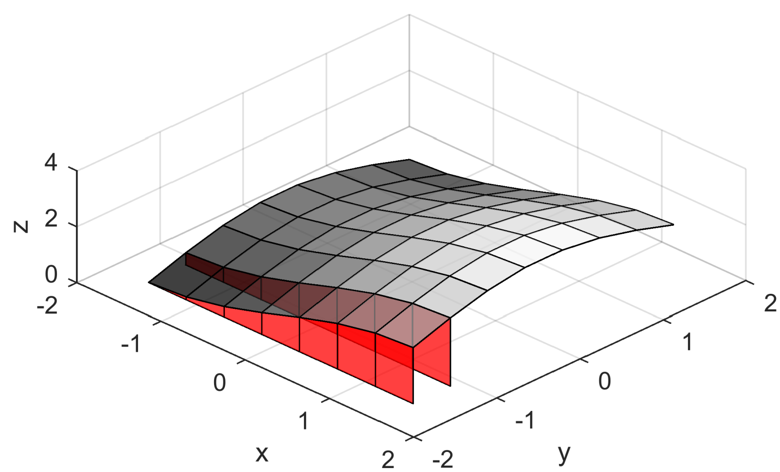

Figure 2

Second, the volume between the two slices is evaluated using the trapezoidal rule:

$$V_1 = \frac{A_1 + A_2}{2} \left((-\pi/2 + \pi/5) - (-\pi/2) \right)$$

Summing up all the sub-volumes V_j defined by slices in the y-direction, we can determine the total volume as follows:

$$V = \sum_j V_j = \sum_j w_j A_j = \sum_j \sum_i w_j w_i z(x_i, y_j)$$

which is the same as Eq. 2.

c) Extend your program to evaluate V using the Simpson's rule for the 1D integrals.

d) Compute the absolute error of the 2D integral approximation, using 20 sub-intervals in each direction (x,y) for both rules (trapezoidal and Simpson's). Repeat your evaluations for an increasing number of intervals in each direction and plot the error vs. the number of sub-intervals in linear and logarithmic scale (error convergence plot). From the logarithmic plot find the order of accuracy of each integration rule.

Question 2: Monte-Carlo integration

The volume of a d-dimensional ball with radius R is defined as:

$$V_d = \int_{-R}^{R} \cdots \int_{-R}^{R} \chi_d(\boldsymbol{x}) \; d\boldsymbol{x} \tag{3}$$

where

$$\chi_d(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \|\boldsymbol{x}\|_2 \leq R \\ 0, & \text{otherwise} \end{cases}$$

$\boldsymbol{x}$ represents the orthogonal axes in the Cartesian coordinate system (i.e., $\boldsymbol{x} = (x_1, x_2, x_3, \cdots x_d)$). $\chi_d(\boldsymbol{x})$ is the characteristic function, which is 1 inside the ball and 0 outside.

The exact value of the volume in d-dimensions is:

$$V_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)} R^d$$

where Γ is the gamma function and satisifies $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$, and $\Gamma(x+1) = x\Gamma(x)$.

For d=2, the ball is a circle and if $R = 1$ it is a unit circle. The volume of a circle is its area. Eq. 3 for d=2 becomes:

$$\begin{aligned} V_2 &= \int_{y=-1}^{1} \int_{x=-1}^{1} \chi_2(x,y) \; dx \; dy \\ \chi_2(x,y) &= \begin{cases} 1, & \text{if } \sqrt{x^2+y^2} \leq 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

This integral can be numerically computed with the Monte-Carlo (MC) method. The method is analogous to 'throwing darts' at a square (Fig. 3). The area of the circle is approximately the ratio of darts that fall within the circle multiplied by the area of the square. When $d = 3$,

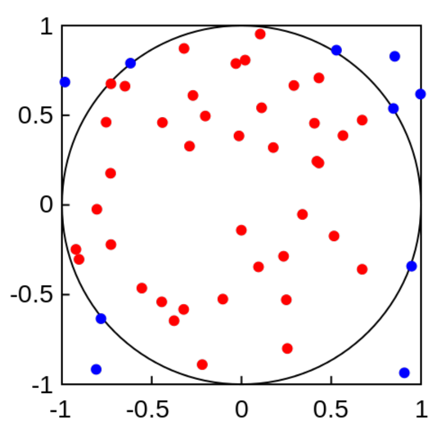

Figure 3: Random points used in MC integration to determine the area of the circle.

the same method involves selecting random points inside a cuboid to determine the volume of a sphere (3d-ball).

a) Compute the volume of a unit ball in 3 dimensions (i.e., a 3D sphere), by evaluating the integral shown in Eq. 3 using the trapezoidal rule with the Cartesian-product approach, as in Question 1. Use 30 uniform intervals in each dimension.

b) Write a program to compute the volume of a 3D unit ball using MC integration, up to the same error as that obtained using the trapezoidal rule in part (a). How many computations do you need? The error for the MC integral for the 3D setting is defined:

$$\varepsilon = V\sqrt{\frac{\langle \chi_3^2 \rangle - \langle \chi_3 \rangle^2}{M}},$$

where $\langle \ldots \rangle$ denotes the average of MC samples, so that $\langle \chi_3 \rangle$ is the average normalised integral of χ_3, V is the volume of the 3D-square and M is the number of sample points.

c) How does the error of the MC approximation scale with M? If you perform MC integration for M = 100000 and you want to reduce the error by 4, how many samples M do you need?

d) Plot the error in the volume of the 3D unit ball computed with MC integration for M varying from 1000 to 10^6. Confirm from your plot that the error scales as $\epsilon \sim 1/\sqrt{M}$.

e) Compute the volume of a 10-dimensional unit ball using MC integration to a relative error of 0.01. How many computations do you need for this?

f) One of the biggest advantages of MC integration is that the error does not change significantly with increasing dimensionality (d) of the problem. Determine if this is indeed true, by plotting error against d, for $d = 2$ to $d = 10$, using $M = 10^6$.