Exam Problem 11

Andreas Nygaard

June 21, 2020

Problem Description

Multidimensional pseudo-random (plain Monte Carlo) vs quasi-random (Halton and/or lattice sequence) integrators

Investigate the convergence rates (of some interesting integrals in different dimensions) as function of the number of sample points.

Solution

I have implemented two sequences for quasi-random Monte Carlo integration, which is the Lattice sequence and the Halton sequence. These low-discrepancy sequences should converge faster than the pseudo-random Monte Carlo integration, since we make sure that the sampling points are more evenly distributed. Because the sampling points of the pseudo-random method are statistically independent, we have a finite probability of all sampling points ending up in only one half of the integration region. This independency is eliminated in the quasi-random case, which ensures a faster convergence [1]. The convergence rate (the decrease of error) of the pseudo-random method is $\mathcal{O}(1/\sqrt{N})$, where N is the number of sampling points, and the convergence rate of the quasi-random method is close to $\mathcal{O}(1/N)$. For an integral of dimension s we can put an upper limit on the convergence rate at $\mathcal{O}(\log(N)^s/N)$, which means that the error will at least decrease as fast as this [2]. We will see that it actually converges faster than this upper bound, but not quite as fast as the $\mathcal{O}(1/N)$ estimate either.

Using 5000 sampling points, I have calculated the integral of the Himmelblau's function, $f(x,y) = (x^2+y-11)^2+(x+y^2-7)^2$, from 0 to 6 in both dimensions, which analytically is

$$\int_0^6 \mathrm{d}x \int_0^6 \mathrm{d}y (x^2+y-11)^2+(x+y^2-7)^2 = 11\,390.4 \tag{1}$$

In table 1 we see the results of a calculation with these three methods. The estimated errors of the quasi-random methods are no good, since the central limit theorem doesn't apply when the points are not statistically independent [1]. Instead we use the difference in the integral estimates of the two quasi-random sequences Lattice and Halton as an error estimate in the following treatment.

	Pseudo-Random	Lattice	Halton
Integral	11 268.07	11 296.72	11 363.40
Exact Error	122.33	93.68	27.00
Error Estimate	175.17	-	-

Table 1: *Estimates of the integral of the Himmelblau's function using all three sampling methods. All of these calculations have* $N = 5000$.

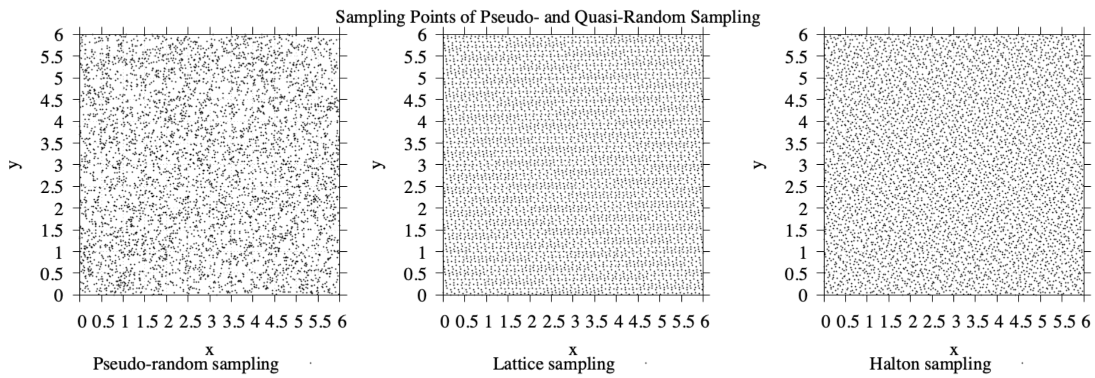

Figure 1: *Sampling points of the three methods (Pseudo-random, Lattice sequence and Halton sequence) when calculating the integral of the Himmelblau's function. The number of points is* $N = 5000$ *for all three methods.*

Figure 1 shows the sampling points of the three methods when calculating the integral of the Himmelblau's function. The lattice is very clear in the Lattice sequence, and the Halton sequence gives a much more evenly distributed sampling than the pseudo-random Monte Carlo.

We will use the Himmelblau's function to test the error as a function of N. The result of this is seen in figure 2, where we show both the estimated and actual error of both the pseudo-random and quasi-random method. The actual error of the quasi-random method is measured from the estimated integral of the Lattice sequence method, while the estimated error is the difference between the estimated integrals of Lattice and Halton as mentioned previously. The figure also shows the $\mathcal{O}(1/\sqrt{N})$ behavior separately along with $\mathcal{O}(\log(N)^s/N)$ behaviors for $s = 0, 1, 2$.

We will also check the convergence of a 5D integral from 0 to 2 in alle five dimensons:

$$\int_{[0,2]^5} \mathrm{d}V f(x,y,z,p,q) = \int_0^2 \mathrm{d}x \int_0^2 \mathrm{d}y \int_0^2 \mathrm{d}z \int_0^2 \mathrm{d}p \int_0^2 \mathrm{d}q x \cdot p^2 - 10z \cdot q + 5y^2 = -64 \tag{2}$$

The result is seen in figure 3, where we show the same as before but now with $s = 0, 1, 2, 3, 4, 5$. A zoom-in on the figure also shows the behavior of the convergence at relatively small N.

We clearly see that the error of the quasi-random method falls faster with respect to N than that of the pseudo-random method which falls as $\mathcal{O}(1/\sqrt{N})$, both in 2D and 5D. As mentioned before, an upper bound on the error estimation should fall as $\mathcal{O}(\log(N)^s/N)$, where s is the dimension. In the

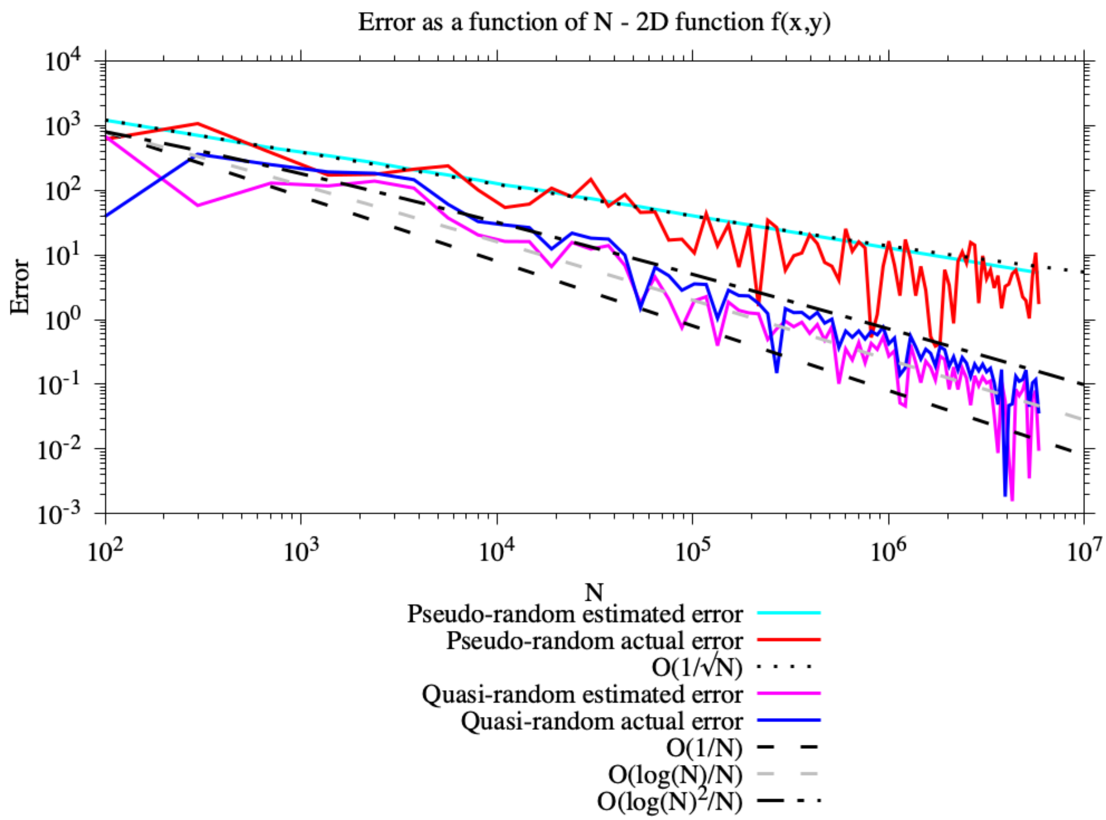

Figure 2: *Pseudo- and quasi-random convergence rates for a 2D integral. The integral is that of the Himmelblau's function:* $f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$, *from 0 to 6 in both dimensions.*

2D case, it is not clear from the figure whether the quasi-random error falls as $\mathcal{O}(\log(N)^2/N)$ or as $\mathcal{O}(\log(N)/N)$, but we can however see that it doesn't fall slower than $\mathcal{O}(\log(N)^2/N)$ and it doesn't fall as fast as just $\mathcal{O}(1/N)$. When examining the 5D case, we see that the $\mathcal{O}(\log(N)^5/N)$ behavior might hold for smaller N (still large enough to give a reasonable result though), but for very large N it seems to fall down to an $\mathcal{O}(\log(N)^2/N)$ behavior or at least faster than the upper bound. We can once again think of the $\mathcal{O}(1/N)$ behavior as a lower bound, which brings the conclusion:

For quasi-random methods, the error as a function of N falls somewhere between the behaviors $\mathcal{O}(1/N)$ (lower bound) and $\mathcal{O}(\log(N)^s/N)$ (upper bound) where s is the dimension, thus making it faster converging than the pseudo-random method with convergence rate $\mathcal{O}(1/\sqrt{N})$.

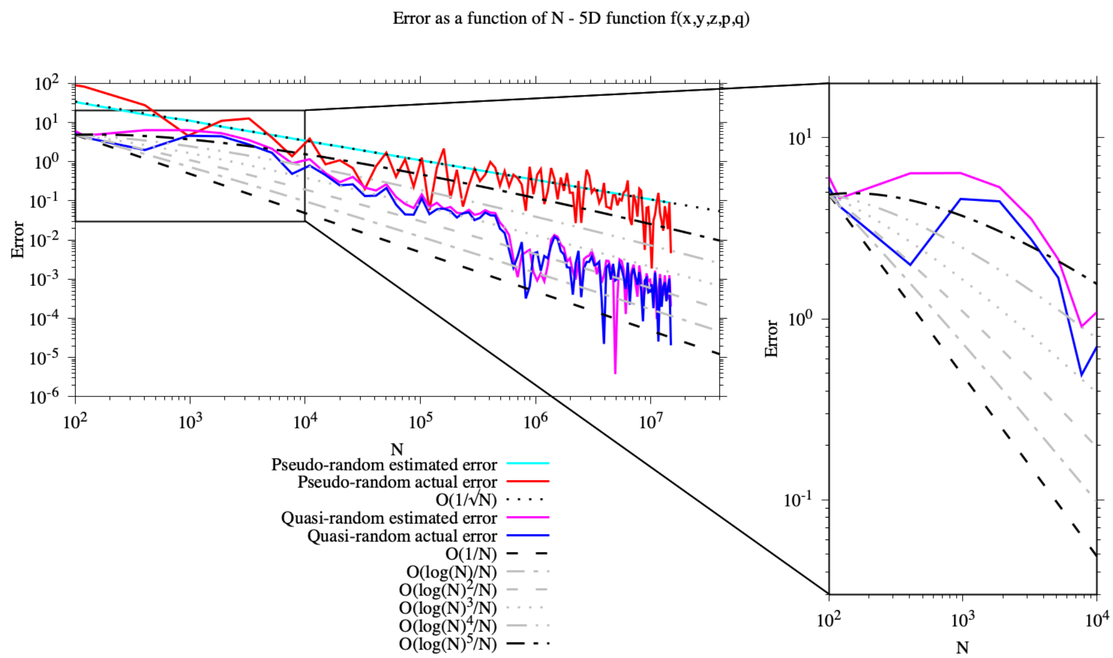

Figure 3: *Pseudo- and quasi-random convergence rates for a 5D integral. The integral is that of the function:* $f(x, y, z, p, q) = x \cdot p^2 - 10z \cdot q + 5y^2$, *from 0 to 2 in all five dimensions.*

References

[1] D. V. Federov, *Monte Carlo integration*, available at http://62.107.14.89/~fedorov/prog/book/montecarlo.pdf.

[2] *Quasi-Monte Carlo method*, available at http://en.wikipedia.org/wiki/Quasi-Monte_Carlo_method.