Continuous Maths HT 2019: Problem Sheet 3

Integration in Higher Dimensions, Convergence, 1-Dimensional Root-Finding

3.1 The two-dimensional Trapezium rule, with n strips in each dimension, is

$$\int_{[a,b]\times[c,d]} f(x,y)\,\mathrm{d}(x,y) \approx \sum_{i=0}^{n}\sum_{j=0}^{n} w_{ij} f\Big(a + i\big(\tfrac{b-a}{n}\big), c + j\big(\tfrac{d-c}{n}\big)\Big)$$

where $\mathbf{W} = (w_{ij})$ is a $(n+1)\times(n+1)$ matrix of scalars that depend only on n, a, b, c, d. By iterating the one-dimensional Trapezium rule for $\int_a^b \int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x$, derive the matrix $\mathbf{W}$. What would you expect in three dimensions?

3.2 Most of the numerical integration methods we have seen are not, as we have described them, iterative. But Monte Carlo integration can be made iterative: each iteration adds one random sample and updates the estimate. Describe $MC_{N+1}[f,R]$ in terms of $MC_N[f,R]$ and, measuring error by the estimate's standard deviation, prove that this iterative algorithm has logarithmic convergence.

Hint: the following limit may be helpful: as $n \to \infty$, $q(n) = \frac{n^{-1/2}-(n+1)^{-1/2}}{n^{-3/2}} \to \frac{1}{2}$.

3.3 Suppose that we wish to approximate $\int_R f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ by partitioning R into k predetermined disjoint sub-regions $R_1,\ldots,R_k$ and using Monte Carlo integration on each sub-region:

$$\int_R f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_i \int_{R_i} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \approx \sum_i MC_{N_i}[f,R_i]. \qquad (*)$$

We wish to find the best $N_1,\ldots,N_k$ with $\sum_i N_i = N$. For the purposes of this question we will ignore the fact that each N_i must be an integer, finding the real-valued N_i that give the most accurate estimate, and assuming that rounding to integers makes negligible difference.

Find the variance of the estimate $(*)$, in terms of the N_i, $A(R_i)$, and $\mathrm{Var}[f(\boldsymbol{X_i})]$ where $\boldsymbol{X_i}$ is generated uniformly in R_i. Show that finding the optimal allocation of samples between sub-regions is related to the constrained optimization problem

$$\underset{N_1,\ldots,N_k}{\text{minimize}} \sum_i \tfrac{a_i}{N_i} \text{ subject to } \textstyle\sum_i N_i = N,\ \forall i.\, N_i \geq 0.$$

Solve this, and suggest an algorithm for Monte Carlo integration given f, the sub-regions $R_1,\ldots,R_n$, and for each i the value of $\mathrm{Var}[f(\boldsymbol{X_i})]$ where $\boldsymbol{X_i}$ is generated uniformly in R_i.

For a challenge, suggest an algorithm for the case when the variances $\mathrm{Var}[f(\boldsymbol{X_i})]$ *are unknown.*

3.4 Consider the recurrence relation $x_{k+2} = \frac{9}{4}x_{k+1} - \frac{1}{2}x_k$.

(a) Find the general solution. Verify that with initial conditions $x_1 = \frac{1}{3}$ and $x_2 = \frac{1}{12}$ the solution is $x_k = \frac{4}{3}\big(\frac{1}{4}\big)^k$.

(b) Implement the recurrence relation on a computer, and use it to graph or tabulate x_k for $k = 1, 2, \ldots, 100$. What happened? Why?

3.5 For a constant $L > 0$, the iteration

$$x_{n+1} = x_n + Le^{-x_n} - 1$$

is an example of Newton's method to find a zero of a function $f : \mathbb{R} \to \mathbb{R}$.

(a) What was $f(x)$, and what is the root x^*? What could the iteration be used for?

(b) Determine an interval for x_0 that guarantees quadratic convergence.

(c) It is a fact that this iteration converges for any x_0, but not quickly. Without running the iteration, estimate approximately how many steps it will take to converge, with relative error at most 10^{-10}, if $L = 2$ and $x_0 = 100$. Then if $L = 2$ and $x_0 = -100$.

(d) Given a computer which cannot calculate logarithms, suggest sensible choices for ___ in the following combination of algorithms to find x^* with relative error at most 10^{-10}:

'Initialize $(a_0, b_0) =$ ___. Run interval bisection until ___. Then initialize $x_0 =$ ___ and run Newton's method until ___. The approximate root is ___.'

3.6 (a) Let $f : [a, b] \to \mathbb{R}$ be convex, differentiable, and strictly decreasing, with $f(b) < 0$. Let $x_0 \in [a, b)$ satisfy $f(x_0) \geq 0$. Show that Newton's method, starting from x_0, always produces an increasing sequence $(x_0, x_1, \ldots)$ that converges to a root of f.

(In fact, Newton's method converges for all *convex functions, and all concave functions too, even if they are not monotone, as long f has a root where its derivative is not zero.)*

Now consider **Example 5.4** from the lectures, where we wanted to find the extinction probability for a Poisson(λ) branching process, $\lambda > 1$. This is the *lower* root of

$$f(x) = e^{\lambda(x-1)} - x = 0. \tag{$\dagger$}$$

We cannot apply Newton's method blindly, because we want to avoid converging to the other root of ($\dagger$), $x = 1$, and we also need to avoid the point where $\frac{\mathrm{d}f}{\mathrm{d}x} = 0$.

(b) Compute $\frac{\mathrm{d}f}{\mathrm{d}x}$ and $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}$; find the unique x^+ such that $\frac{\mathrm{d}f}{\mathrm{d}x}(x^+) = 0$. Verify that a root of ($\dagger$) is in $(0, x^+)$, and that the conditions of part (a) apply to f on this interval.

(c) Suggest how to apply Newton's method to find the extinction probability. Write a program to implement this, and tabulate or graph the probability as λ varies.

3.7 In this question we will analyze the convergence of the Secant method. Let x^* be a root of $f(x) = 0$ where $\frac{\mathrm{d}f}{\mathrm{d}x}(x^*) \neq 0$ and $\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(x^*) \neq 0$, and write $\epsilon_n = x_n - x^*$.

(a) Using the Secant method iteration, Lemma 5.6 from the lecture notes, and Taylor's theorem, show that

$$\epsilon_{n+1} = \epsilon_n\epsilon_{n-1}\left(\frac{\frac{f(x_n)}{\epsilon_n} - \frac{f(x_{n-1})}{\epsilon_{n-1}}}{\epsilon_n - \epsilon_{n-1}}\right)\left(\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right) = \tfrac{1}{2}\epsilon_n\epsilon_{n-1}\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(\beta)\Big/\frac{\mathrm{d}f}{\mathrm{d}x}(\alpha),$$

for some α, β in an interval that includes x_{n-1}, x_n, and x^*.

(b) Show that the Secant method converges to x^* as long as both x_0 and x_1 lie in $(x^* - c, x^* + c)$, and the same conditions hold as in Lemma 5.2(ii).

(c) Explain (without giving a formal proof) why $\frac{|\epsilon_{n+1}|}{|\epsilon_n\epsilon_{n-1}|} \to C$, for some $C > 0$.

(d) Find the order of convergence, by supposing that $\frac{|\epsilon_{n+1}|}{|\epsilon_n|^q} \to a$, then solving for q and a.