ECON 698R
Problem Set 4
Due Monday, April 9

John Rust

Spring 2002

I. The *discrepancy* of a set of n points $(t_1,\ldots,t_n)$ in $[0,1]^d$ (the d-dimensional hypercube) is given by

$$D_n^*(t_1,\ldots,t_n) = \sup_{B\in\mathcal{B}} |\lambda_n(B) - \lambda(B)|, \tag{1}$$

where $\mathcal{B}$ is the set of all *normalized subrectangles* of $[0,1]^d$,

$$\mathcal{B} = \left\{ B \subset [0,1]^d | B = \prod_{i=1}^d [0,b_i], \quad b_i \in [0,1] \right\}, \tag{2}$$

and $\lambda(B)$ is the *Lebesgue measure* of B,

$$\lambda(B) = \prod_{i=1}^d b_i, \tag{3}$$

and $\lambda_n(B)$ is the *empirical measure* of B,

$$\lambda_n(B) = \frac{1}{n}\sum_{i=1}^n I\{s_i \in B\}. \tag{4}$$

where $I\{s\in B\}$ is the indicator function,

$$I\{s\in B\} = \begin{cases} 1 & \text{if} \quad s\in B \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

1. Consider the one dimensional case, $d = 1$. Find a formula for the discrepancy, $D_n^*(t_1,\ldots,t_n)$.
2. Find a formula for the *minimal discrepancy* points $(t_1^*,\ldots,t_n^*)$, i.e. the points that solve

$$(t_1^*,\ldots,t_n^*) = \underset{(t_1,\ldots,t_n)}{argmin} D_n^*(t_1,\ldots,t_n). \tag{6}$$

II. Consider the problem of deriving the optimal (deterministic) algorithm that has the smallest worst case error for integrating the Lipschitz continuous functions on $[0,1]$ with uniform Lipschitz bound L. The problem asks you to fill out the details in the derivation sketched in chapter 1 of J.F. Traub and A.G. Werschulz (1998) *Complexity and Information* (Cambridge University Press). They consider integration in the worst case setting over the class of functions $\mathcal{F}$ satisfying

$$|f(x) - f(x)| \le L|x-y|, \quad x,y\in[0,1], \quad f\in\mathcal{F}. \tag{7}$$

Recall that an *algorithm* for computing the integral of a function $f\in\mathcal{F}$ can be written as a composition of two functions, $\phi_n : R^n \to R$ and $I_n : [0,1]^n \to R^n$ where $I_n = (f(s_1),\ldots,f(s_n))$ is the *information* (or sample) on the unknown function f we wish to integrate and ϕ_n is an algorithm that combines this information into an estimate of the integral:

$$\int_0^1 f(x)dx \approx \phi_n(f(s_1),\ldots,f(s_n)). \tag{8}$$

Thus, we are seeking a rule for choosing the sample points $(s_1,\ldots,s_n)$ and the function ϕ_n that minimize the worst case integration error:

$$r(n) = \inf_{s_1,\ldots,s_n} \inf_{\phi_n} \sup_{g\in\mathcal{F}(f(s_1),\ldots,f(s_n))} \left|\phi_n(f(s_1),\ldots,f(s_n)) - \int_0^1 g(x)dx\right|, \tag{9}$$

where

$$\mathcal{F}(f(s_1),\ldots,f(s_n)) = \{g\in\mathcal{F}|g(s_1) = f(s_1),\ldots,g(s_n) = f(s_n)\}. \tag{10}$$

Thus, $\mathcal{F}(f(s_1),\ldots,f(s_n))$ is the equivalence class of functions in $\mathcal{F}$ that have the same information (i.e. have the same values over the n sample points $(s_1,\ldots,s_n)$) as the true function f that we are trying to integrate. Since we assume that we don't know the true f at all points but only at the n points $(s_1,\ldots,s_n)$, we consider the worst case error by computing the function $g\in\mathcal{F}(f(s_1),\ldots,f(s_n))$ whose actual integral is as far away as possible from the approximate integral $\phi_n(f(s_1),\ldots,f(s_n))$.

1. Show that $\mathcal{F}(f(s_1),\ldots,f(s_n))$ is a set of functions in $\mathcal{F}$ bounded above by an *upper envelope* $\overline{f}$ and and a *lower envelope* $\underline{f}$, and that $\overline{f}$ and $\underline{f}$ are piecewise-linear functions with slopes everywhere equal to $\pm L$ that satisfy

$$\underline{f}(s_i) = \overline{f}(s_i) = f(s_i), \quad i = 1,\ldots,n. \tag{11}$$

2. Show that the optimal algorithm ϕ_n^* is given by

$$\phi_n^*(f(s_1),\ldots,f(s_n)) = \int_0^1 f_{\text{mid}}(x)dx, \tag{12}$$

   where $f_{\text{mid}} = (\underline{f}+\overline{f})/2$.

3. Let f_{pwl} be the piecewise linear interpolant of the points $(0,f(s_1)),(s_1,f(s_1)),\ldots,(s_n,f(s_n)),(1,f(s_n))$. Show that

$$\int_0^1 f_{\text{mid}}(x)dx = \int_0^1 f_{\text{pwl}}(x)dx. \tag{13}$$

4. Show that

$$\int_0^1 f_{\text{pwl}}(x)dx = f(s_1)s_1 + \sum_{i=1}^{n-1}\frac{1}{2}(f(s_i)+f(s_{i+1})(s_{i+1}-s_i) + f(s_n)(1-s_n). \tag{14}$$

   Thus, the optimal integration algorithm for the class $\mathcal{F}$ is the *modified trapezoidal rule.*

5. Show that there is no loss of generality in restricting attention to the special case of *zero information*, i.e. where $f(s_1) = f(s_2) = \cdots = f(s_n) = 0$, i.e. show that

$$\begin{aligned}\sup_{g\in\mathcal{F}(f(s_1),\ldots,f(s_n))}\left|\phi_n^*(f(s_1),\ldots,f(s_n)) - \int_0^1 g(x)dx\right| &= \sup_{g\in\mathcal{F}(0,\ldots,0)}\left|\phi_n^*(0,\ldots,0) - \int_0^1 g(x)dx\right| \\ &= \sup_{g\in\mathcal{F}(0,\ldots,0)}\left|\int_0^1 g(x)dx\right|. \end{aligned}\tag{15}$$

6. Show that
$$\sup_{g\in\mathcal{F}(0,\ldots,0)}\left|\int_0^1 g(x)dx\right| = L\left(\frac{1}{2}s_1^2+\frac{1}{4}\sum_{i=1}^{n-1}(s_{i+1}-s_i)^2+\frac{1}{2}(1-s_n)^2\right). \tag{16}$$

7. Using calculus, derive the optimal placement of the sample points. Show that the optimal points satisfy
$$s_i^* = \frac{2i-1}{2n} \tag{17}$$
show that the worst case error bound using the optimal integration algorithm and the optimally placed points satisfies
$$r(n) = \frac{L}{4n}. \tag{18}$$

8. Show that for the optimally chosen points, the optimal integration algorithm takes the form of a *quasi monte carlo algorithm:*
$$\phi_n^*(f(s_1^*),\ldots,f(s_n^*)) = \frac{1}{n}\sum_{i=1}^{n} f(s_i^*). \tag{19}$$
How does this result compare with the quasi monte carlo algorithm for the minimum discrepancy points $(t_1^*,\ldots,t_n^*)$ derived in problem I-1?