1 Bounds for Bracketing Methods

1.1 Curvature Bound Gives Function Bound

Given a twice-differentiable function $f(x)$ such that:

- f has a global minimum at x^*
- For all $x < x^*, f'(x) < 0$ and for $x > x^*, f'(x) > 0$
- $0 < f''(x) < \dfrac{1}{K}$

and any interval (a, b) that brackets x^* such that $\delta \stackrel{\text{def}}{=} b - a \leq K$:

Theorem 1.1. *For all* $x \in (a, b)$, $f(x) - f(x^*) < K - \sqrt{K^2 - \delta^2}$.

Proof. Salient facts:

- Osculating circles.
- Squeeze theorem.

□

1.2 Function Bound Gives Algorithm Step Bound

Given:

- $f(x)$ as before on a domain of length m, with $0 < f''(x) < \dfrac{1}{K}$
- ϵ vertical tolerance

Theorem 1.2. *The golden section algorithm will achieve* ϵ *tolerance at iteration*

$$n = \frac{\log\left(K^2 - (K - \epsilon)^2\right)}{\log \varphi} - \frac{\log m}{\log \varphi}$$

where $\varphi = \dfrac{\sqrt{5} - 1}{2}$

Proof. The golden section algorithm by construction reduces the interval width from m to φm ($\varphi \stackrel{\text{def}}{=} \dfrac{\sqrt{5}-1}{2}$) every iteration. From there, use the prior result and the fact that $\varphi^n \to 0$ as $n \to \infty$. □

2 Stochastic Golden Section Search

2.1 Algorithm

With the function f as before and the domain $[0, m]$; define the function $\hat{f}(x) = f(x) + Z_x$ where Z_x is a stochastic noise term at x. We assume that any distinct Z_x noise terms are *independent and normal*. The following procedure is an adaptation of golden section search to find the minimum of $f(x)$ given only the observations $\hat{f}(x)$.

Define the *candidate interval* $[a, b]$. Initialize the interval with $a = 0$ and $b = m$. Choose a significance level $1 - \alpha$, first-stage sample count $n_0 \geq 2$, and a precision target ϵ.

- Test Point Selection: Let $x_1 = a + (1 - \varphi)(b - a)$ and $x_2 = a + \varphi(b - a)$.
  - Initialize: Take n_0 first-stage samples from $\hat{f}(x_1)$ and $\hat{f}(x_2)$. Construct the estimates for the mean $\bar{X}_1$ and $\bar{X}_2$ and variance S_1 and S_2.
  - Stopping: If $\left|\bar{X}_1 - \bar{X}_2\right| \geq c\left(\frac{S_1}{\sqrt{n_1}} + \frac{S_2}{\sqrt{n_2}}\right)$ (where c is given from the choice of α based on the normal distribution) then select either the interval $[a, x_2]$ or $[x_1, b]$ based on which of $\bar{X}_1$ or $\bar{X}_2$ is smaller. Otherwise, proceed to refinement.
  - Refinement: If $S_1\sqrt{(n_2 + 1)n_2}\left(\sqrt{n_1 + 1} - \sqrt{n_1}\right) \geq S_2\sqrt{(n_1 + 1)n_1}\left(\sqrt{n_2 + 1} - \sqrt{n_2}\right)$ then sample once from $\hat{f}(x_1)$ and update $n_1, \bar{X}_1, S_1$ accordingly; otherwise, sample once from $\hat{f}(x_2)$ and update similarly. After updating, check the stopping rule.
  - Escape: If $\max\left(c\frac{S_1}{\sqrt{n_1}}, c\frac{S_2}{\sqrt{n_2}}\right) < \epsilon$ then the precision target has been reached- select $[a, x_2]$ as the interval.
- Set the *candidate interval* $[a, b]$ to the selected interval.
- Stopping: If $b - a < \sqrt{K^2 - (K - \epsilon)^2}$, theorem 1.1 implies that the precision target has been attained.

2.2 Analysis

Given a function with curvature bound $0 < f''(x) < \frac{1}{K}$ and domain $[0, m]$, the algorithm with precision target ϵ will require $n(K, m, \epsilon)$ iterations as given from theorem 1.2. From this it is possible to examine from a finite sample perspective the probability of correct selection and the expected error. The process requires each confidence interval to correctly cover the true mean of the function $\hat{f}(x)$; with $n(K, m, \epsilon)$ iterations this requires $n(K, m, \epsilon) + 1$ confidence intervals to be correct, placing the lower bound on probability of correct selection at $(1 - \alpha)^{n(K,m,\epsilon)+1}$.

The worst-case expected error can be approached by considering the worst-case scenario where the true solution is at one extreme of the interval. From this, we can set up the following recursive formula for the error at stage $k+1$, E_{k+1}:

$$E_{k+1} = \begin{cases} E_k & \text{correct selection} \\ E_k + (1-\varphi)\varphi^k m & \text{incorrect selection} \end{cases}$$

Applying the law of total expectation, we have the following recursive formula for $\mathbb{E}(E_k)$:

$$\mathbb{E}(E_{k+1}) = \alpha\mathbb{E}(E_k) + (1-\alpha)\left(\mathbb{E}(E_k) + (1-\varphi)\varphi^k m\right)$$

$$\mathbb{E}(E_{k+1}) = \mathbb{E}(E_k) + \alpha(1-\varphi)\varphi^k m$$

Solving this recurrence relation yields

$$\mathbb{E}(E_k) = \alpha\varphi(1-\varphi^k)m$$

Thus, after $n(K, m, \epsilon)$ iterations, the expected error is $\alpha\varphi(1-\varphi^{n(K,m,\epsilon)})m$.

3 Stochastic Bisection With Golden Section

3.1 Problem

Given a function $g(x, d)$ where d is a discrete variable on $[1, N]$ and x is a continuous variable on $[0, m]$. Suppose that for any fixed d, the function $g_d(x) \stackrel{\text{def}}{=} g(x, d)$ satisfies the assumptions as given before in 1.1, with the curvature bound $0 < \frac{\partial^2}{\partial x^2} g_d(x) < \frac{1}{K_d}$. Additionally, suppose that for any fixed x the function $g(x, d)$ is monotonically decreasing. Finally, we suppose only noisy observations of $g(x, d)$ may be made. To solve the following constrained optimization problem (c is some fixed constraint), a mixture of bisection and golden section algorithms are applied.

$$\min d$$

$$\exists x : g(x, d) \leq C$$

3.2 Algorithm

Choose a tolerance level ϵ, and significance level α. Set the *discrete candidate interval* $[N_1, N_2]$ to $[1, N]$.

- Perform the stochastic golden section algorithm with ϵ and α on the function $g_d(x)$ where $d = N_1 + \lfloor \frac{N_2 - N_1}{2} \rfloor$ with the additional stopping criterion of stopping if the confidence interval of either test point lies fully below the constraint level C.

- If the golden section algorithm has found a point x where the confidence interval for $g_d(x)$ is below the constraint then set the discrete candidate interval to $[N_1, d]$ and repeat; otherwise, set the discrete candidate interval to $[d, N_2]$ and repeat.

- Once $N_2 - N_1 = 1$, stop and claim N_2 is the solution with the correct value of x having been covered in the golden section process.

Optionally, an additional error checking procedure can be introduced to the bisection search based on backtracking; once $N_2 - N_1 = 1$:

- Perform the golden section algorithm on both $g_{N_1}(x)$ and $g_{N_2}(x)$ until the minimum is found.
- If the minima values are both below C, then set $N_2 = N_1$ and set $N_1 = N_1 - 1$ and repeat.
- If the minima values are both above C, then set $N_1 = N_2$ and set $N_2 = N_2 + 1$ and repeat.
- Once the minima for $g_{N_1}(x)$ lies above C and the minima for $g_{N_2}(x)$ lies below C, stop and claim that N_2 is the solution.

3.3 Analysis

3.3.1 Without Error Checking

Once again an analysis of the finite sample behavior is possible. With N total possible values for the discrete value, then the bisection algorithm will terminate within $\log_2\lceil N\rceil$ steps; hence, if the probability of correct selection for each golden section search process is $(1-\beta_d)$ let $(1-\beta) \stackrel{\text{def}}{=} \min_{d\in[1,N]}(1-\beta_d)$. The probability of correct selection for the binary search is thus bounded below by $(1-\beta)^{\log_2\lceil N\rceil}$.

The expected error for the binary search procedure can be bounded above by considering the worst case. Suppose there are $2^M = N$ total indices to check (for a non power of two, throwing in as many "dummy" cases as necessary to attain a power of two will suffice). Suppose there is a constant probability of incorrect selection β at every stage. In the worst case, the true solution is at one extreme of the interval; suppose without loss of generality that the true solution is at index 1. Now, it is already known that the bisection algorithm will take M steps to complete.

Once again a recurrence relation and the law of total expectation lead to the result. In this case, the recurrence relation for error at step k is:

$$E_{k+1} = \begin{cases} E_k & \text{correct selection} \\ E_k + \dfrac{N}{2^{k+1}} & \text{incorrect selection} \end{cases}$$

This results in the expected error at step k in closed form as

$$\mathbb{E}(E_k) = \frac{N\beta}{2^k}(2^k - 1)$$

3.3.2 With Error Checking

Pending...

Future thoughts:

- If the bisection step is positive, consider starting the next process with a shortened continuous candidate interval (due to the monotonic nature); be wary of erroneous selection causing problems, however.
- More intelligent probability of correct selection- not assuming a constant PCS across all of it for ex