GOLDEN SECTION METHOD

What is it?

Unconstrained Optimization Algorithms

- Direct Search
- Line Search
  - Initial point
  - Search direction
  - $\boxed{\text{Distance to move along search direction}}$ ↗ find α^*
  - When to stop
- Trust Region

gives us a way to search along a line

- An approach that <u>refines the bracketing of the minimum</u> in order to converge to an estimate of the minimum

- Takes its name from the "golden section" or "golden ratio":

$$\boxed{\frac{1+\sqrt{5}}{2} \approx 1.61803}$$

What do we need to know? (assumptions)

- Function is <u>unimodal</u> along the search direction
- $\left.\frac{df}{d\alpha}\right|_1 < 0$, i.e. the <u>search is in a descent direction</u>

} a minimum defined by α^* and f^* exists along the search direction

<u>Unimodal Function</u> – contains only one minimum or maximum on the interval in question

Advantages?

- <u>Robust method</u>

Disadvantages?

- <u>Can take a long time</u> (can require a lot of function calls)

$$N = \frac{\ln \varepsilon}{\ln(1-\tau)} + 3,$$

N: number of function calls

ε: tolerance

τ: reduction in interval size ($\tau = 0.38197$)

number of function calls needed to achieve a particular tolerance

$$\varepsilon = (1-\tau)^{N-3}$$

"-3" accounts for the function calls in the initial step

Main Points:

Golden Section Method <u>defines a distance</u> to move <u>along a search direction</u> (refine the bracketing of the minimum) <u>based on the "golden ratio."</u> It is a <u>robust method</u> (reduces bounds by same amount after every iteration), but it <u>can take a long time</u> (requires a lot of function calls).

GOLDEN SECTION METHOD

How does it work?

1. You are given two points (α_U and α_L) and the function values at those points ($f(\alpha_U)$ and $f(\alpha_L)$).

$\underset{\uparrow \atop \text{lower}}{\alpha_L} < \underset{\uparrow \atop \text{upper}}{\alpha_U}$, the minimum is bracketed

Where do we place a_1 and a_2?

2. Now we need two intermediate points between α_L and α_U (which we will call α_1 and α_2).

between α_U and α_L

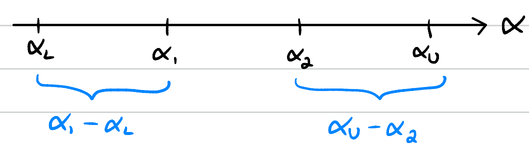
"The most efficient algorithm is the one that reduces the bounds by the same fraction on each iteration." – Vanderplaats

can be formally proven

if we "throw away" either α_U or α_L and add a new point, we would "throw away" the same fraction of the original interval $(\alpha_U - \alpha_L)$ each time

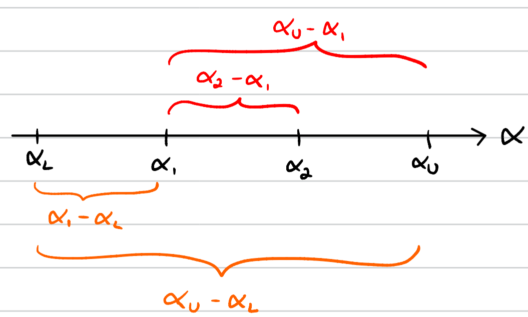
Let's start by making the problem symmetric:

$$\alpha_1 - \alpha_L = \alpha_U - \alpha_2$$



Now we make a requirement based on reducing the bounds by the same amount after each iteration:

$$\frac{\alpha_1 - \alpha_L}{\alpha_U - \alpha_L} = \frac{\alpha_2 - \alpha_1}{\alpha_U - \alpha_1}$$



What we are saying so far:

If we "throw away" α_L above, we are requiring that

$$\underset{\text{new interval width}}{\overset{\text{new difference in two lowest bounds}}{}} \quad \boxed{\frac{\alpha_2 - \alpha_1}{\alpha_U - \alpha_1} = \frac{\alpha_1 - \alpha_L}{\alpha_U - \alpha_L}} \quad \underset{\text{old interval width}}{\overset{\text{old difference between two lowest bounds}}{}}$$

By making this requirement, we will always choose α such that we eliminate the same fraction of the interval every time

GOLDEN SECTION METHOD

Method Requirements:

- Symmetry: $\alpha_1 - \alpha_L = \alpha_U - \alpha_2$ => no matter if α_U or α_L is discarded, the results will be the same

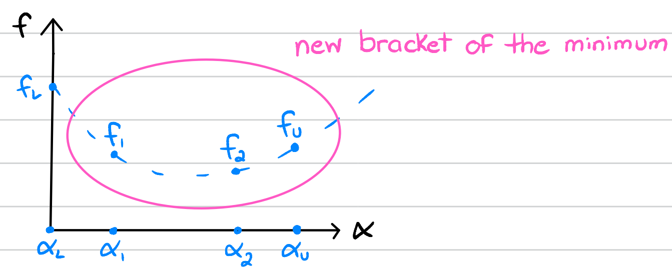
- Same fractional reduction of the interval after each iteration: $\dfrac{\alpha_1 - \alpha_L}{\alpha_U - \alpha_L} = \dfrac{\alpha_2 - \alpha_1}{\alpha_U - \alpha_1}$

(We can say the same thing if we discard α_U instead: $\dfrac{\alpha_U - \alpha_2}{\alpha_U - \alpha_L} = \dfrac{\alpha_2 - \alpha_1}{\alpha_1 - \alpha_L}$)

These requirements allow us to <u>specify</u> an appropriate <u>spacing rule</u> to <u>define</u> <u>specific locations along the α axis</u> to <u>add new points</u> (α_1, α_2, etc.)

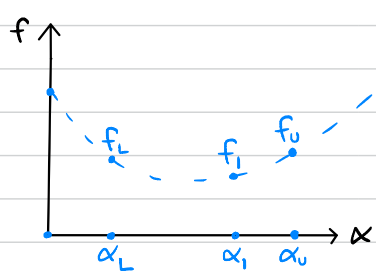
How does it work? (continued)

3. Now adding α_1 and α_2:



In this case, $f_2 < \min(f_U, f_1)$ so α_2 refined the bracket => discard α_L

4. Relabel the points by shifting the labels and repeat the process



(In a similar manner, if $f_1 < \min(f_L, f_2)$, discard α_U => α_2 becomes α_U, α_1 becomes α_2)

What are the values of α_1 and α_2?

1. Assume $\alpha_L = 0$ and $\alpha_U = 1$

2. $\dfrac{\alpha_1 - \alpha_L}{\alpha_U - \alpha_L} = \dfrac{\alpha_2 - \alpha_1}{\alpha_U - \alpha_1}$ => $\alpha_1 = \dfrac{\alpha_2 - \alpha_1}{1 - \alpha_1}$

But we have the requirement that $\alpha_1 - \alpha_L = \alpha_U - \alpha_2$ => $\alpha_1 = 1 - \alpha_2$ or $\alpha_2 = 1 - \alpha_1$

Main points:

If $f(\alpha_1) \leq \min\{f(\alpha_L), f(\alpha_2)\}$
1. Discard α_U
2. Change α_2 to α_U
3. Change α_1 to α_2

If $f(\alpha_2) \leq \min\{f(\alpha_U), f(\alpha_1)\}$
1. Discard α_L
2. Change α_1 to α_L
3. Change α_2 to α_1

GOLDEN SECTION METHOD

What are the values of α_1 and α_2?

Plugging this in:

$$\alpha_1 = \frac{(1-\alpha_1)-\alpha_1}{1-\alpha_1} = \frac{1-2\alpha_1}{1-\alpha_1}$$

Doing some algebra: $\alpha_1(1-\alpha_1) = \alpha_1 - \alpha_1^2 = 1-2\alpha_1$

$$0 = \alpha_1^2 - \alpha_1 - 2\alpha_1 + 1$$
$$0 = \alpha_1^2 - 3\alpha_1 + 1$$

Solving for α_1: $\alpha_1 = \frac{3 \pm \sqrt{(-3)^2 - 4(1)(1)}}{2(1)} = \frac{3 \pm \sqrt{5}}{2} \approx 0.38197, 2.61803$

2.61803: outside the original interval of [0, 1] ⇓ meaningless

thus, $\alpha_1 \approx 0.38197$

$$\alpha_2 = 1 - \alpha_1 \approx 0.61803$$

and the ratio of $\frac{\alpha_2}{\alpha_1}$ gives us the "golden ratio": $\boxed{\frac{\alpha_2}{\alpha_1} \approx \frac{0.61803}{0.38197} \approx 1.61803}$

Normalized bounds:

If we let $\tau = 0.38197$, then we can write:

$$\boxed{\begin{aligned} \alpha_1 &= (1-\tau)\alpha_L + \tau\alpha_U \\ \alpha_2 &= \tau\alpha_L + (1-\tau)\alpha_U \end{aligned}}$$

Convergence and function calls:

Consider a relative tolerance, ε, defined as: $\varepsilon = \frac{\Delta\alpha}{\alpha_{U,0} - \alpha_{L,0}}$

where $\alpha_{U,0}$: initial upper bound $\alpha_{L,0}$: initial lower bound

$\Delta\alpha = \alpha_U - \alpha_L$ at each iteration

Because the interval is reduced by $\tau = 0.38197$ each iteration,

$$\varepsilon = (1-\tau)^{N-3}$$ where N: total number of function calls

$$\Rightarrow N = \frac{\ln\varepsilon}{\ln(1-\tau)} + 3$$

← we can calculate the number of function calls to achieve a particular tolerance

← we can estimate the relative reduction in the bounds around the minimum for a specified number of function calls

Main Points:

We found $\alpha_1 = 0.38197$ and $\alpha_2 = 0.61803$ by setting $\alpha_L = 0$ and $\alpha_U = 1$

$$\boxed{\begin{aligned} \alpha_1 &= (1-\tau)\alpha_L + \tau\alpha_U \\ \alpha_2 &= \tau\alpha_L + (1-\tau)\alpha_U \end{aligned}}$$

These equations give us α_1 and α_2 for any value of α_L and α_U

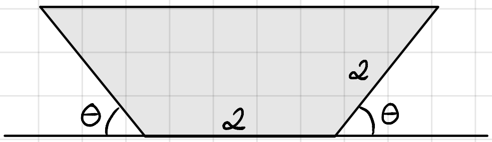
GOLDEN SECTION METHOD

Example:

Consider the figure below. The cross-sectional area A of a gutter with equal base and edge length of 2 is given by:

$$A = 4\sin\theta(1+\cos\theta)$$

Find the angle θ which maximizes the cross-sectional area of the gutter. Using an initial interval of $[0, \frac{\pi}{2}]$, find the solution after 2 iterations. Use an initial $\varepsilon = 0.05$.



Solution:

The function to be maximized is $f(\theta) = 4\sin\theta(1+\cos\theta)$

1. Calculate the values of α_1 and α_2

$$\alpha_1 = (1-\tau)\alpha_L + \tau\alpha_U$$
$$\alpha_2 = \tau\alpha_L + (1-\tau)\alpha_U$$

where $\tau = 0.38197$

$$\alpha_1 = (1-0.38197)(0) + (0.38197)\left(\tfrac{\pi}{2}\right) = 0.6$$
$$\alpha_2 = (0.38197)(0) + (1-0.38197)\left(\tfrac{\pi}{2}\right) = 0.9708$$

2. Calculate f_L, f_1, f_2, f_U

$$\left.\begin{aligned} f_L(0) &= 4\sin(0)(1+\cos(0)) = 0 \\ f_1(0.6) &= 4\sin(0.6)(1+\cos(0.6)) = 4.1227 \\ f_2(0.9708) &= 4\sin(0.9708)(1+\cos(0.9708)) = 5.1654 \\ f_U\left(\tfrac{\pi}{2}\right) &= 4\sin\left(\tfrac{\pi}{2}\right)\left(1+\cos\left(\tfrac{\pi}{2}\right)\right) = 4 \end{aligned}\right\} \quad f(\alpha_2) \geq \max\{f(\alpha_1), f(\alpha_U)\}$$

$\Rightarrow$ discard α_L

$$\left.\begin{aligned} &\alpha_1 \text{ becomes } \alpha_L \\ &\alpha_2 \text{ becomes } \alpha_1 \end{aligned}\right\} \begin{aligned} \alpha_L &= 0.6 \\ \alpha_1 &= 0.9708 \\ \alpha_U &= \tfrac{\pi}{2} \end{aligned}$$

3. Calculate new intermediate value

$$\alpha_2 = \tau\alpha_L + (1-\tau)\alpha_U = (0.38197)(0.6) + (1-0.38197)\left(\tfrac{\pi}{2}\right) = 1.2$$

$$\frac{\alpha_2 - \alpha_L}{\alpha_1 - \alpha_L} = \frac{1.2-0.6}{0.9708-0.6} = 1.61812 \checkmark$$

4. Check the stopping criteria

$\alpha_U - \alpha_L = \frac{\pi}{2} - 0.6 = 0.9708 > 0.05 \Rightarrow$ repeat the process

5. Calculate f_L, f_1, f_2, f_U

$$\left.\begin{aligned} f_L(0.6) &= 4\sin(0.6)(1+\cos(0.6)) = 4.1227 \\ f_1(0.9708) &= 4\sin(0.9708)(1+\cos(0.9708)) = 5.1654 \\ f_2(1.2) &= 4\sin(1.2)(1+\cos(1.2)) = 5.0791 \\ f_U\left(\tfrac{\pi}{2}\right) &= 4\sin\left(\tfrac{\pi}{2}\right)\left(1+\cos\left(\tfrac{\pi}{2}\right)\right) = 4 \end{aligned}\right\} \quad f(\alpha_1) \geq \max\{f(\alpha_2), f(\alpha_L)\}$$

$\Rightarrow$ discard α_U

α_2 becomes α_U

α_1 becomes α_2

$$\left.\begin{matrix} \\ \\ \end{matrix}\right\} \begin{aligned} \alpha_L &= 0.6 \\ \alpha_2 &= 0.9708 \\ \alpha_U &= 1.2 \end{aligned}$$

6. Calculate new intermediate value

$$\alpha_1 = (1-\tau)\alpha_L + \tau\alpha_U = (1-0.38197)(0.6) + (0.38197)(1.2) = 0.82918$$

$$\frac{\alpha_2-\alpha_L}{\alpha_1-\alpha_L} = \frac{0.9708-0.6}{0.82918-0.6} = 1.61794 \checkmark$$

7. Check the stopping criteria

$$\alpha_U - \alpha_L = 1.2 - 0.6 = 0.6 > 0.05 \Rightarrow \text{Repeat the process}$$

Thus, at the end of the second iteration, the solution is:

$$\frac{\alpha_U + \alpha_L}{2} = \frac{1.2+0.6}{2} = 0.9$$ ← the average of the upper and lower boundary points

Therefore, the maximum area occurs when $\boxed{\theta = 0.9 \text{ radians or } 51.6^\circ}$

Note that the search would stop after the 9^{th} iteration ($\varepsilon < 0.05$) where the optimum value is calculated as $\theta = 1.0416$ (59.68°).

The theoretical optimal solution to the problem happens exactly at 1.0472 radians (60°)