Geometric Mean for Trapezoidal Integration

William John Holden

10 April 2021

I had the idea to use the geometric mean to estimate a numerical integral using the trapezoid rule. In this informal paper, I will non-rigorously explore the idea with Julia.

The area under the continuous function f between a and b can be estimated by the trapezoid rule:

$$(b-a)\frac{f(a)+f(b)}{2}$$

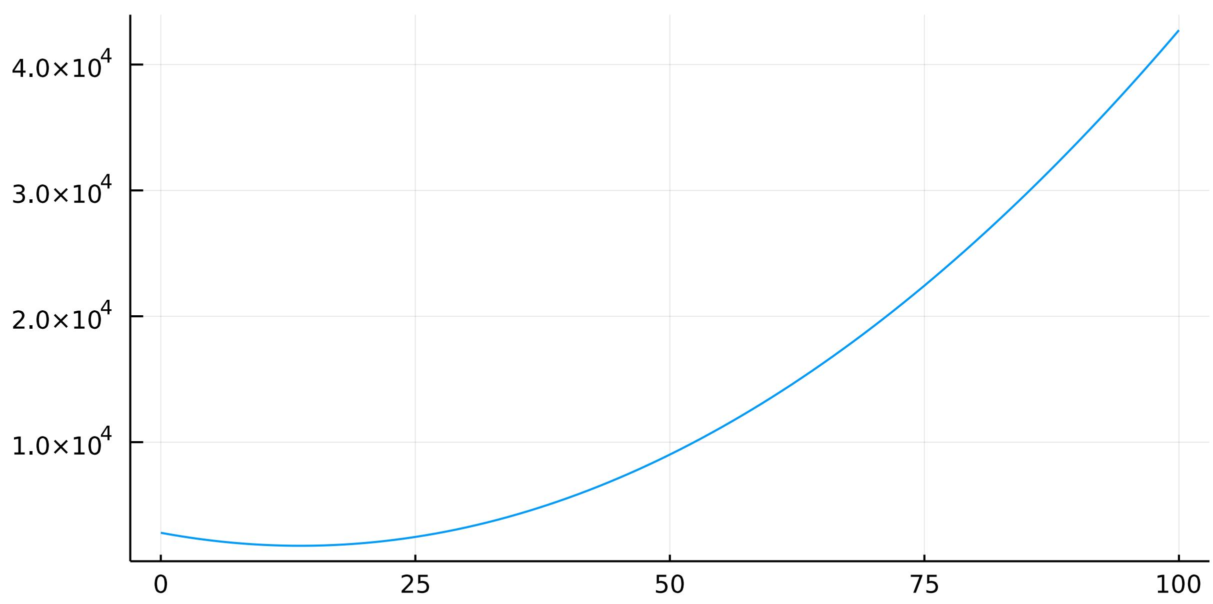
The value $(f(a)+f(b))/2$ is the average height of a trapezoid with width $b-a$. This formula produces an exact result if f is a linear equation or constant. Otherwise, this formula contains some error, ε. A lower value for ε means a better estimate.

$$\int_a^b f(t)\,dt = (b-a)\frac{f(a)+f(b)}{2} + \varepsilon$$

Particularly, if $f \in \Theta(n^2)$, then perhaps the geometric mean, $\sqrt{ab}$, could give a better estimate for the integral.

Let $f(t)$ be a quadratic function.

```
f(t) = 5.5 * t^2 - 150.7 * t + 2800;
using Plots;
plot(0:100, f, size=(600,300), legend=false)
```



The antiderivative, $F(t)$, is easily computed by the power rule.

```
F(t) = 5.5 / 3 * t^3 - 150.7 / 2 * t^2 + 2800 * t;
```

We estimate the integral of an arbitrary function over a specified interval using a user-supplied averaging function.

```
function integrate(fcn, interval, average)
    y = map(fcn, interval)
    dx = step(interval)
    return sum([dx * average(y[i], y[i + 1]) for i in 1:(length(y) - 1)])
end;
```

We estimate

$$\int_{-10}^{10} f(t)\,dt$$

using a named range from -10 to 10 in 0.1-unit intervals.

```
x = -10:.1:10;
```

Using the arithmetic mean, we compute the numerical integral as:

```
arithmetic(fcn, interval) = integrate(fcn, interval, (a , b) -> (a + b) / 2);
arithmetic(f, x)
```

```
## 59666.85000000001
```

The geometric mean gives a similar estimate. (Note that this function is delicate and error-prone with negative values).

```
geometric(fcn, interval) = integrate(fcn, interval, (a, b) -> sqrt(a * b));
geometric(f, x)
```

```
## 59666.654799561846
```

The definite (true) integral integral is

```
definite(Fcn, x) = Fcn(last(x)) - Fcn(first(x));
definite(F, x)
```

```
## 59666.666666666664
```

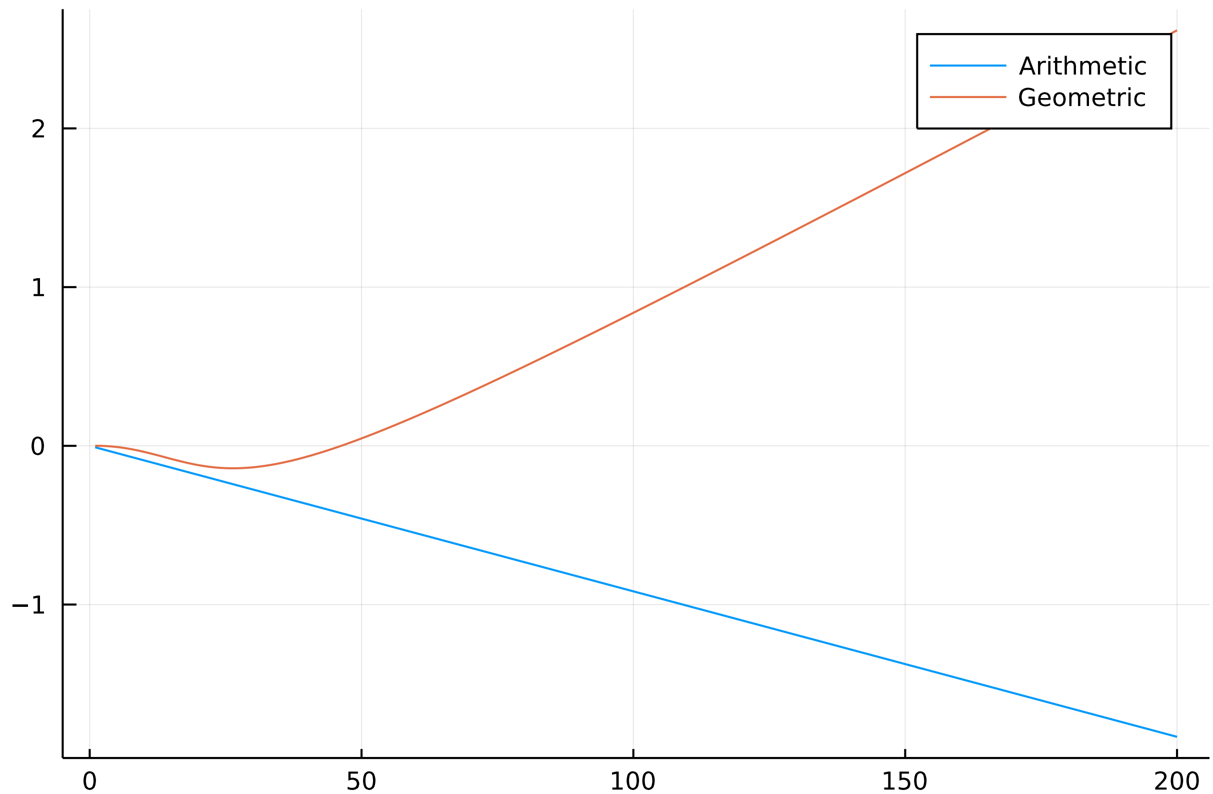
As shown, both give a good estimate. The geometric mean has slightly less error, but is this always the case? Let's see how each estimator performs with the same function but a different interval.

```
using DataFrames
errors = DataFrame(
  :Arithmetic => [definite(F, 0:.1:x2) - arithmetic(f, 0:.1:x2) for x2 in 1:200],
  :Geometric => [definite(F, 0:.1:x2) - geometric(f, 0:.1:x2) for x2 in 1:200])
```

```
## 200×2 DataFrame
##  Row   Arithmetic   Geometric
##        Float64      Float64
##
##    1   -0.00916667   0.00049958
##    2   -0.0183333    3.62675e-5
##    3   -0.0275      -0.0014109
##    4   -0.0366667   -0.00385249
##    5   -0.0458333   -0.00728602
```

```
##    6   -0.055        -0.0116934
##    7   -0.0641667    -0.0170382
##    8   -0.0733333    -0.023264
##
##  194   -1.77833       2.50941
##  195   -1.7875        2.52747
##  196   -1.79667       2.54554
##  197   -1.80583       2.56361
##  198   -1.815         2.58169
##  199   -1.82417       2.59977
##  200   -1.83333       2.61785
##                  185 rows omitted
```

```
plot([errors.Arithmetic, errors.Geometric], label=["Arithmetic" "Geometric"])
```



The above scatter plot shows the width of the interval, $x = b - a$, against the error of the integral estimate $y = \varepsilon$. The plot shows that on the intervals starting at 0 and ending at $0 < x \leq 100$, the geometric mean gives the better estimate. Interestingly, somewhere near $x = 45$ or so, the error is nearly zero. As x exceeds 100, though, the geometric mean underestimates the integral more than the arithmetic mean overestimates it. As with interval width continues to grow, errors accumulate more rapidly using the geometrical mean than the arithmetic mean.

It looks like the geometric mean is *not* as interesting for the trapezoid rule as I had speculated. Under some circumstances, it *may* produce a better estimate, but with some loss in generality. Even with a quadratic function, which seems like an ideal case for this approach, the geometric mean gives a less accurate estimate than the arithmetic mean.