The Babylonian Method and Its Properties

Gersham Johnson

Abstract: The computation of square roots is one of the oldest algorithmic procedures in mathematics. The Babylonian method is one of the oldest methods of root approximation, and it is still a valuable tool for understanding both the development and functioning of root-taking methods. We will examine the precision and effectiveness of this method, and compare it to a few other well-known methods. We will also establish the validity of this method by proving its convergence. Finally, we will seek to observe any trends that may lead to a more generalized version.

I. Introduction

Root computations were among the first algorithmic techniques to be developed in mathematics, and the Babylonians were among the first to utilize them. The Babylonians established one of the first civilizations in human history, and by about 1800 B.C. they had already developed a significant and influential body of mathematics. Though they in fact had no explicit algorithms, the Babylonians were the first to derive the algorithmic procedure that bears their name. It would take many centuries before it was explicitly stated by Heron of Alexandria (Heath 340), but the method is a relatively simple iterative process based on a simple rule. To begin, the square root of a number A is a number x such that $x^2 = A$. For **perfect squares**, whose roots are integers, computations of roots are simple. However, for other real numbers the task becomes more difficult; approximation techniques are needed to derive a value.

II. The Method

In order to approximate $\sqrt{A}$, begin with an initial approximation, $x_1 = x$. From this value a second approximation, $\frac{A}{x}$, can be established such that $x \cdot \frac{A}{x} \approx A$. It necessarily follows that by averaging the two values, a second and better approximation is achieved. Let this approximation equal x_2. This method can be understood by evaluating it with cases:

Case 1: $x = \sqrt{A}$. The new approximation thus becomes $\frac{\sqrt{A}+\frac{A}{\sqrt{A}}}{2} = \frac{2\sqrt{A}}{2} = \sqrt{A}$, and we are done.

Case 2: $x < \sqrt{A}$. With this approximation, $\frac{A}{x} > \sqrt{A}$, and thus $x < \sqrt{A} < \frac{A}{x}$. The new approximation $\frac{x+\frac{A}{x}}{2}$ will thus be closer than either x or $\frac{A}{x}$.

Case 3: $x > \sqrt{A}$. For this case, $\frac{A}{x} < \sqrt{A}$, and the reasoning is analogous.

We have thus established the method set forth by the Babylonians, but we have not in fact established its validity. In what manner does the series of iterations converge, if at all? For this we turn to the Ancient Greeks, who recognized that the arithmetic mean of two numbers a, b is always greater than or equal to the geometric mean: $\frac{a+b}{2} \geq \sqrt{ab}$. So the new approximation x_2 will always be greater (or equal) than $\sqrt{A}$ (and **only** equal when a=b). But what of each successive approximation? Taking x_2 and applying the method again we get $x_3 = \frac{x_2+\frac{A}{x_2}}{2}$. As we have shown, $x_2 \geq \sqrt{A}$, so $\frac{A}{x_2} \leq \sqrt{A}$. Since $\frac{A}{x_2} \leq x_2$, we have $\frac{A}{x_2} + x_2 \leq x_2 + x_2$, which implies that $\frac{x_2+\frac{A}{x_2}}{2} = x_3 \leq x_2$. Continuing this, we see that, for a given n (really, sufficiently large n: the initial $\frac{A}{x}$ value may be greater than x_1), $x_{n+1} \leq x_n$. Likewise, for sufficiently large n, $\frac{A}{x_{n+1}} \geq \frac{A}{x_n}$. (Both inequalities can be proved with induction.) So let $\{a_n\}$ be the sequence of terms of the form $\frac{A}{x_n}$ and let $\{b_n\}$ be the sequence of terms of the form x_n. Ignoring the first few terms (which do not affect the convergence of a sequence), we see that $\{a_n\}$ is monotonically increasing and $\{b_n\}$ is monotonically decreasing. Equally important is the fact that $\{a_n\}$ is bounded above by $\sqrt{A}$ and $\{b_n\}$ is bounded below by $\sqrt{A}$. Since all bounded and monotonic sequences converge, it is clear that $\{a_n\}$ and $\{b_n\}$ converge ($\{a_n\}$ to its least upper bound, $\{b_n\}$ to its greatest lower bound)–and therefore so does the

Babylonian method. What remains to be seen is the *value* to which it converges. Let's begin by evaluating the difference between x_{n+1} and $\frac{A}{x_{n+1}}$. $\left|x_{n+1} - \frac{A}{x_{n+1}}\right| =$

$\left|\frac{x+\frac{A}{x}}{2} - \frac{2A}{x+\frac{A}{x}}\right| = \left|\frac{x^2+A}{2x} - \frac{2Ax}{x^2+A}\right| = \left|\frac{(x^2-A)^2}{2x(x^2+A)}\right| = \left|\frac{(x^2-A)}{x^2+A}\right| \cdot \left|\frac{(x^2-A)}{2x}\right|$, where $\left|\frac{(x^2-A)}{x^2+A}\right| \leq 1$. Thus $\left|\frac{x+\frac{A}{x}}{2} - \frac{2A}{x+\frac{A}{x}}\right| \leq \left|\frac{(x^2-A)}{2x}\right| = \frac{x+\frac{A}{x}}{2}$

. So $\left|x_{n+1} - \frac{A}{x_{n+1}}\right| \leq \frac{1}{2} \cdot \left|x_n - \frac{A}{x_n}\right|$. With induction we can see that this relationship remains true as $n \to \infty$. Since the difference between the two approximations approaches 0 (the difference decreases by greater than or equal to $\frac{1}{2}$ at each iteration), their limits must therefore be the same.

Now, we have established that $\{a_n\}$ and $\{b_n\}$ are bounded above and below, respectively, and that they both converge. However, take note that $\{a_n\}$ converges to its *least* upper bound. Now, suppose that the least upper bound is some number $L < \sqrt{A}$. Then $\{a_n\}$ must necessarily converge to L. But recall that $\sqrt{A}$ $\leq \{b_n\}$. This means that the greatest lower bound K of $\{b_n\}$ is greater than L , so $\lim_{n\to\infty} \{a_n\} < \lim_{n\to\infty} \{b_n\}$. But this is a contradiction, as we know from above. So $\sqrt{A}$ must be the least upper bound, so both sequences, and the Babylonian method, must therefore converge to the square root of A.

Convergence should not be very surprising. In general, the Babylonian method converges relatively quickly. For $\sqrt{3}$, using the estimate $x = 1.5$ we have $\frac{1.5+\frac{3}{1.5}}{2} = 1.75$, which is already about 0.02 away from the actual value $\sqrt{3} \approx 1.73$. In the next section we will attempt to establish trends relating to the convergence of the method.

III. Convergence of Method

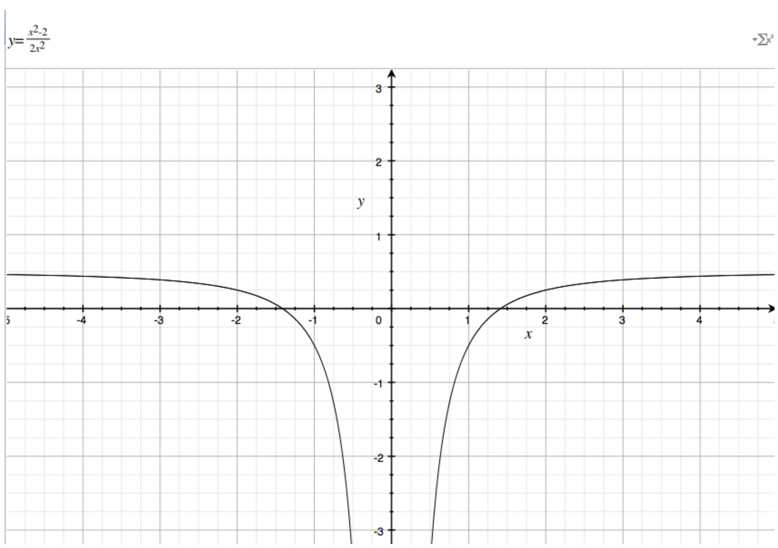
We have so far proved that the Babylonian method does in fact converge, but exactly how fast does it converge? In order to investigate this question we must first establish the significance of the rate of convergence.

Let's begin by defining a function f for the Babylonian method such that $f(x) = \frac{x+\frac{A}{x}}{2}$.

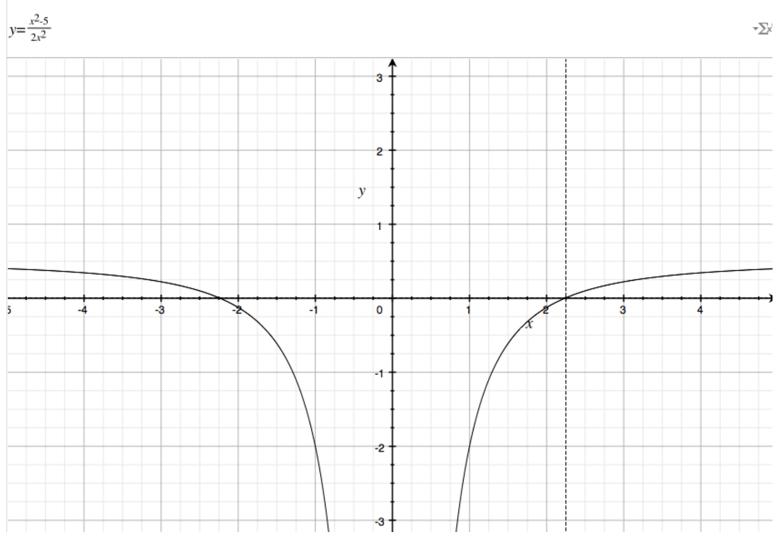
The Babylonian method is of course iterative but observing the rate of change of this function at very good estimates of $\sqrt{A}$ may allow us to see how the function converges at certain seeds.

For this root approximation function, we can observe the rate of change at any point of approximation by taking the derivative of f at that point. With our root approximation function, we have $f'(x) = \frac{x^2-Ax}{2x^2}$. Observing this graph for a series of functions, a few conclusions can be drawn:

$\sqrt{2}$:



$\sqrt{5}$:



For both graphs, as is the case with any graph of the derivative of the Babylonian function at a given root, the location of the root is precisely at $f'(x) = 0$, as simple algebra can verify. This of course provides a convenient way to graphically estimate square roots, but it also poses some questions. One significant issue is the relation between the derivative function and the rate of convergence of the actual iterative method.

Certainly, since the seed $\sqrt{A}$ will continually generate itself in the iterative process, the rate of change is 0. $\sqrt{A}$ is known as a fixed point, and it is in fact very important in establishing convergence. However, the behavior of the derivative function *around* the fixed point is much different. Using $\sqrt{8}$ as our model, we see that the number of iterations required for varying seeds to converge is not directly linked to the derivative, but is *modelled* by its behavior. For an initial seed of $x = 2$, approximately 4 iterations are required to produce an approximation accurate to 7 decimal places. Since $f'(2) = -\frac{1}{2}$, it would be expected that another seed of varying rate of change would converge either more quickly or more slowly.

Let $x = 4$. Then $f'(4) = \frac{1}{4}$. Since the (absolute) rate of change is smaller (which is validated by the fact that 4 is a greater distance away from $\sqrt{8}$ than 2), it would be expected that the rate of convergence would be greater. In fact, $f'(4) = \frac{|f'(2)|}{2}$, so one could expect the rate of convergence to be about twice as big. As it turns out, this is not the case. The rate of convergence for the seed $x = 4$ is the *exact* same as that of the seed $x = 2$. In spite of the differing rates of change of the *initial* approximation, the actual value of the initial approximations, $f(2)$ and $f(4)$, are equal–

$\frac{2+\frac{8}{2}}{2} = \frac{4+\frac{8}{4}}{2} = 3$. So the rate of change of this Babylonian approximation function is not in fact directly indicative of the rate of convergence. However, it may be implicitly related. As $x \to \infty$, the rate of convergence increases, and $\lim_{x\to\infty}(\frac{x^2-A}{2x^2}) = \frac{1}{2}$. However, this limit does not explain why the increasingly negative slopes as x approaches 0 (such as at $x = 0.5$) do not result in a slower rate of convergence. In order to attempt to explain this we must recognize the fact that negative seeds converge to $-\sqrt{A}$, not the positive root. In this case we see that negative seeds converge the same as their positive counterparts. That is, the seeds x and $-x$ converge to $\sqrt{A}$ and $-\sqrt{A}$, respectively.

Table for $\sqrt{8}$:

	X_n	X_n			
1	2	4	5.1	25	0.5
2	3	3	3.3343137	12.66	8.25
3	2.8333333	2.8333333	2.8668040	6.6459557	4.6098484
4	2.8284313	2.8284313	2.8286839	3.9248476	3.1726317
5	**2.8284271**	**2.8284271**	**2.8284271**	2.9815716	2.8470988
6	2.8284271	2.8284271	2.8284271	2.8323601	2.8284883
7	2.8284271	2.8284271	2.8284271	2.8284298	**2.8284271**
8	2.8284271	2.8284271	2.8284271	**2.8284271**	

The Babylonian method is an iterative method, so a variety of important properties can be established. Though we have established that the Babylonian method will converge, we can now establish the fact that it will converge for *all* seeds x. First we must establish the *fixed point* of the iterated function. A fixed point is a value that continually cycles to itself as it is iterated–that is, it is a value for which $f(x) = x$. By setting $\frac{x+\frac{A}{x}}{2} = x$, we simplify to get $x^2 - A = 0$. Thus the fixed point is $\sqrt{A}$. Using the derivative function we have defined, we can evaluate it at this fixed point to determine if it is an *attractive* fixed point. If the fixed point is attractive, every seed will converge to it.

Let's start by evaluating f' at $\sqrt{A}$: $f'(\sqrt{A}) = \frac{(\sqrt{A})^2 - A}{2(\sqrt{A})^2} = 0$. For any iterative function with a fixed point x, if $|f'(x)| < 1$, the fixed point is attractive. As we have shown, $\left|f'(\sqrt{A})\right| = 0 < 1$. So all (non-zero) seeds will converge to the fixed point. And, as can be proved with the definition of the convergence of an iterative method[1], the Babylonian method converges quadratically (the number of correct decimal places approximately doubles at each iteration).

For any value we choose, we can approximate the square root. But the question still remains: is there a way to determine the relative speeds of certain seeds, and which seeds are better than others? As we have seen, seeds that are further from the root converge at a much slower rate. Though it is not apparent how to generalize any results regarding convergence of this iterative method at this point, observing its relation to another method may prove beneficial.

IV. Newton's Method

The Babylonian method is indeed one of the oldest algorithmic procedures, but it is most easily derived from a more modern mathematical technique based in calculus. Newton's method is a general technique for finding the roots of equations, and thus works particularly well for square roots when the given equation is a quadratic of the form $x^2 - A = 0$. With Newton's method, the tangent line to the curve at a given approximation of the root is used to locate a more accurate approximation by finding its own x -intercept. The general method is as follows:

Take an estimate, x_n ,and evaluate it at f and f'. Take the ratio of $f(x_n)$ to $f'(x_n)$ and subtract it from x_n.

[1] See appendix

Note that this is just a rewritten form of the point-slope formula for the equation of the tangent line at x_n solved for the x-intercept:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Newton's method is another example of an iterated function. And using $f(x) = x^2 - A$, a more familiar form surfaces:

$x - \frac{x^2 - A}{2x} = \frac{2x^2 - x^2 + A}{2x} = \frac{x + \frac{A}{x}}{2}$. Thus we can derive the Babylonian method from the similarly efficient Newton's method.

Newton's method also shares other properties with the Babylonian method. For example, both methods display quadratic convergence (for Newton's method this has been proved with the aid of Taylor's theorem), and both methods converge rapidly. However, Newton's method does *not* always converge for all seeds (largely due to badly behaved derivatives.) We know that some seeds do not in fact converge but *diverge*, and the seed $x = 0$ fails (for both methods). The definition of Newton's method with the derivative does enable us to generalize, however. With $f(x) = x^n - A$, we have

$$x_{n+1} = x_n - \frac{x^n - A}{nx^{n-1}} = \frac{nx - x + \frac{A}{x^{n-1}}}{n}.$$

This method of approximation bears a resemblance to various "weighted" versions of the Babylonian method. As educator Dave Elgin observed, a weighted sum utilizing approximations of the form $\frac{A}{x^2}, \frac{A^2}{x^3}, etc.$ enables an approximation of the square root not unlike that of Newton's method (1). In reality, the weighted method is nothing more than a Taylor series approximation:

With N as an approximation to $\sqrt{A}$ (and d as the "error"– $A = N^2 + d$), $\sqrt{N^2 + d} = N + \frac{d}{2N} - \frac{d^2}{8N^3} + \ldots$[2]

A little manipulation reveals that the first two terms are just the Babylonian method! In fact, another

[2] See appendix

ancient algorithmic method–the Bakhshali method–is nothing more than two iterations of the Babylonian method. The reader can verify this themselves:

$$\sqrt{S} \approx N + \frac{d}{2N} - \frac{d^2}{8N^3 + 4Nd}$$

V. Conclusion

We have demonstrated the convergence of the Babylonian method, studied properties of its rate of convergence, and observed a handful of generalizations that have exposed a variety of questions. It has been proved that at least three arithmetical operations are necessary for any root-taking algorithm, but is the Babylonian method truly the most simple algorithm? And what is the true extent of the complexities of Newton's method that are retained in the Babylonian method? We observed various convergence studies but very few claims were established. How can convergence be more accurately measured for algorithms and approximating techniques? The results of this investigation have opened up even more questions and opportunities for problem solving.

It is interesting to note the variety of root approximation methods that "contain" the Babylonian method, from the Taylor Series approximation to Newton's Method. One great area of inquiry would be to find out procedurally how the Babylonians devised their method, and whether or not it bears any resemblance to the other methods highlighted. As can be seen, though the Babylonian method may be one of the most simple and ubiquitous algorithms, its mathematical importance should not be underestimated.

VI. References

Books:

- Heath, Sir Thomas L. "Heron of Alexandria." *A History of Greek Mathematics, Volume II: From Aristarchus to Diophantus*. 2nd ed. Vol. 2. N.p.: Dover Publications, 1981. 298-354. Print.
- Newton, Isaac. *The Method of Fluxions.* London: n.p., 1736. Print.

Journals and Articles:

- Carothers, Neal. "Still More on Square Roots." *Still More on Square Roots*. Bowling Green State University, n.d. Web. 06 Apr. 2013.
- Elgin, David. "Square Root and Cube Root Algorithms." *Mathematics in School* 35.1 (2006): 30-31. *Jstor.org*. The Mathematical Association. Web. 07 Apr. 2013.
- Kosheleva, Olga. "Babylonian Method of Computing the Square Root: Justifications Based on Fuzzy Techniques and on Computational Complexity." *Utep.edu*. The University of Texas at El Paso, n.d. Web. 02 Apr. 2013.
- Osler, Thomas J. "Extending the Babylonian Algorithm." *Rowan.edu*. Rowan University, n.d. Web. 06 Apr. 2013.

VII. Appendix

1. By definition, if $\lim_{k\to\infty} \frac{|x_{k+1}-L|}{|x_k-L|^2} = \mu > 0$, the iterative method converges quadratically. By evaluating the difference between successive approximations, one can verify this for the Babylonian method.

2. In general, a Taylor series centered at $x = c$ for a function f is generated as follows:

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)(x-c)^2}{2!} + ...$$

For $f(x) = \sqrt{x}$, $f(N^2 + d)$ centered at $c = N^2$ yields the Taylor series as seen above. For the first two terms, we say $N + \frac{d}{2N} = \frac{N+\frac{A}{N}}{2}$. Simplifying yields $2(2N^2 + d) = 2N(N + \frac{A}{N})$. This results in $4N^2 + 2d = 2N^2 + 2A$. Dividing by two yields $2N^2 + d = N^2 + A$, and so $A = N^2 + d$, and we are done.