**Title**
Rate of Convergence of the Arithmetic-Geometric Mean Process



**Author**
Leeuw, Jan de

**Publication Date**
2008-04-22

Peer reviewed

# RATE OF CONVERGENCE OF THE ARITHMETIC-GEOMETRIC MEAN PROCESS

JAN DE LEEUW

ABSTRACT. This (didactic) note gives a simple counter-example to the notion that Picard iterations converge super-linearly if and only if the sup-norm of the Jacobian at the solution is equal to zero and sub-linearly if and only if it is equal to one.

## 1. INTRODUCTION

Suppose $S$ is an open subset of $\mathbb{R}^n$ and $\Gamma : S \Rightarrow S$ is a differentiable map. Assume the Picard iterations $x^{(k+1)} = \Gamma(x^{(k)})$ starting from some $x^{(0)} \in S$ converge to $x \in S$. We can derive information about the rate of convergence from the sup-norm (the eigenvalue of maximum modulus) of the derivative $\mathcal{D}\Gamma(x)$. If $\|\mathcal{D}\Gamma(x)\| = \lambda < 1$ we have linear convergence with rate $\lambda$, and if $\|\mathcal{D}\Gamma(x)\| = 0$ we have super-linear convergence [Ortega and Rheinboldt, 1970, Chapter 10]. $\|\mathcal{D}\Gamma(x)\| = 1$ often indicates sub-linear convergence. Our elementary example below, however, has $\|\mathcal{D}\Gamma(x)\| = 1$ and quadratic convergence.

## 2. THE ARITHMETIC-GEOMETRIC MEAN

Suppose $a$ and $b$ are two positive numbers. Their *arithmetic mean* is defined as $\mathbf{AM}(a,b) = \frac{1}{2}(a+b)$ and their *geometric mean* as $\mathbf{GM}(a,b) = \sqrt{ab}$.

*Result* 1. $\mathbf{AM}(a,b) \geq \mathbf{GM}(a,b)$ with equality if and only if $a = b$.

*Proof.* $0 \leq (\sqrt{a} - \sqrt{b})^2 = 2(\mathbf{AM}(a,b) - \mathbf{GM}(a,b))$. □

*Date*: October 3, 2008 — 8h 42min — Typeset in LUCIDA BRIGHT.

From now on suppose, without loss of generality, that $a > b$. Let $a_0 = a$ and $b_0 = b$ and define the sequences

$$a_n = \mathbf{AM}(a_{n-1}, b_{n-1}), \tag{1a}$$

$$b_n = \mathbf{GM}(a_{n-1}, b_{n-1}). \tag{1b}$$

*Result* 2. $a_n > b_n$

*Proof.* From Result 1. □

*Result* 3. $\{a_n\}$ is a decreasing sequence, which is bounded below, and thus converges to some $a_\infty$. $\{b_n\}$ is an increasing sequence, which is bounded above, and thus converges to some $b_\infty$.

*Proof.* $a_n < \mathbf{max}(a_{n-1}, b_{n-1}) = a_{n-1}$ and $b_n > \mathbf{min}(a_{n-1}, b_{n-1}) = b_{n-1}$. Moreover $a_n > b_n > b$ and $b_n < a_n < a$. □

*Result* 4. $a_\infty = b_\infty$.

*Proof.* Take limits on both sides of (1). This gives

$$a_\infty = \mathbf{AM}(a_\infty, b_\infty),$$
$$b_\infty = \mathbf{GM}(a_\infty, b_\infty).$$

Both equations imply $a_\infty = b_\infty$. □

The common limit $a_\infty = b_\infty$ is called the *arithmetic-geometric mean* of $a$ and $b$, written as $\mathbf{AGM}(a, b)$. The arithmetic-geometric mean was studied by Legendre and Gauss, and it has fascinating applications in many areas of mathematics and numerical analysis. There are excellent reviews of these applications in Carlson [1971], Cox [1984], and Almqvist and Berndt [1988].

*Result* 5. b $< \mathbf{GM}(a, b) < \mathbf{AGM}(a, b) < \mathbf{AM}(a, b) < a$

*Proof.* $\mathbf{AGM}(a, b) = a_\infty < a_1 = \mathbf{AM}(a, b) < a_0 = a$ and $\mathbf{AGM}(a, b) = b_\infty > b_1 = \mathbf{GM}(a, b) > b_0 = b$. □

For another proof of the convergence to a common limit we define the sequence $\delta_n = a_n - b_n$. It should be noted that $\delta_n$ is a reasonable way to measure distance to the solution, since

$$|a_n - \mathbf{AGM}(a,b)| + |b_n - \mathbf{AGM}(a,b)| = \\ a_n - \mathbf{AGM}(a,b) + \mathbf{AGM}(a,b) - b_n = \delta_n.$$

*Result* 6. $\{\delta_n\}$ is a decreasing sequence bounded below by zero, and thus converges to some $\delta_\infty \geq 0$.

*Proof.* Since $a_n < a_{n-1}$ and $b_n > b_{n-1}$ we have $\delta_n = a_n - b_n < a_{n-1} - b_{n-1} = \delta_{n-1}$. Moreover $\delta_n > 0$ for all $n$. □

*Result* 7. $\delta_\infty = 0$

*Proof.*

$$\delta_n = \mathbf{AM}(a_{n-1}, b_{n-1}) - \mathbf{GM}(a_{n-1}, b_{n-1}) = \frac{1}{2}(\sqrt{a_{n-1}} - \sqrt{b_{n-1}})^2, \tag{2a}$$

$$\delta_{n-1} = a_{n-1} - b_{n-1} = (\sqrt{a_{n-1}} - \sqrt{b_{n-1}})(\sqrt{a_{n-1}} + \sqrt{b_{n-1}}), \tag{2b}$$

and thus $\delta_n < \frac{1}{2}\delta_{n-1}$ It follows that $0 < \delta_n < (\frac{1}{2})^n \delta_0$ and thus $\lim_{n\to\infty} \delta_n = 0$. □

The proof shows that convergence of $\{\delta_n\}$ is faster than that of a geometric sequence with radius $\frac{1}{2}$. But we can be more precise.

*Result* 8. Convergence of the sequence $\{\delta_n\}$ to zero is superlinear, i.e.

$$\lim_{n\to\infty} \frac{\delta_n}{\delta_{n-1}} = 0.$$

*Proof.* From Equations (2)

$$\frac{\delta_n}{\delta_{n-1}} = \frac{1}{2}\frac{\sqrt{a_{n-1}} - \sqrt{b_{n-1}}}{\sqrt{a_{n-1}} + \sqrt{b_{n-1}}} \to 0.$$

□

In fact, we can be even more precise.

*Result* 9. Convergence of the sequence $\{\delta_n\}$ to zero is quadratic.

$$\lim_{n\to\infty} \frac{\delta_n}{\delta_{n-1}^2} = \frac{1}{8}\frac{1}{\mathbf{AGM}(a,b)}.$$

*Proof.* From Equations (2)

$$\frac{\delta_n}{\delta_{n-1}^2} = \frac{1}{2}\frac{1}{(\sqrt{a_{n-1}}+\sqrt{b_{n-1}})^2} \rightarrow \frac{1}{8}\frac{1}{\mathbf{AGM}(a,b)}.$$

□

In a sense, the sequences $\{a_n\}$ and $\{b_n\}$ converge equally fast.

*Result* 10. $(a_n - a_{n-1}) \sim -(b_n - b_{n-1})$, i.e.

$$\lim_{n\rightarrow\infty}\frac{a_n - a_{n-1}}{b_n - b_{n-1}} = -1.$$

*Proof.*

$$\begin{aligned} a_n - a_{n-1} &= -\frac{1}{2}(\sqrt{a_{n-1}}-\sqrt{b_{n-1}})(\sqrt{a_{n-1}}+\sqrt{b_{n-1}}),\\ b_n - b_{n-1} &= \sqrt{b_{n-1}}(\sqrt{a_{n-1}}-\sqrt{b_{n-1}}). \end{aligned}$$

and thus

$$\frac{a_n - a_{n-1}}{b_n - b_{n-1}} = -\frac{1}{2}\frac{\sqrt{a_{n-1}}+\sqrt{b_{n-1}}}{\sqrt{b_{n-1}}} \rightarrow -1.$$

□

## 3. Counterexample

Equation (1) defines a mapping $\Gamma : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$. The derivative of this mapping is

$$\mathcal{D}\Gamma(a,b) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}\sqrt{\frac{b}{a}} & \frac{1}{2}\sqrt{\frac{a}{b}} \end{bmatrix},$$

and thus

$$\mathcal{D}\Gamma(\mathbf{AGM}(a,b),\mathbf{AGM}(a,b)) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

which has eigenvalues one and zero.

The fact that $\|\mathcal{D}\Gamma(\mathbf{AGM}(a,b),\mathbf{AGM}(a,b))\|_\infty = 1$ seems to suggest sublinear convergence, while in fact we know convergence is quadratic. If $y_n$ is the two-element vector with elements $a_n - a_{n-1}$ and $b_n - b_{n-1}$, normalized to length one, then Result 10 shows that $y_n$ converges to a vector with elements $-1$ and $+1$. This eigenvector corresponds with the smallest eigenvalue of the Jacobian at the solution, and that smallest eigenvalue is equal to zero.

## References

G. Almqvist and B. Berndt. Gauss, Landen, Ramanujan, the Arithmetic-Geometric Mean, Ellipses, $\pi$, and the Ladies Diary. *The American Mathematical Monthly*, 95:585–608, 1988.

B.G. Carlson. Algorithms Involving Arithmetic and Geometric Means. *The American Mathematical Monthly*, 78:496–505, 1971.

D.A. Cox. The Arithmetic-Geometric Mean of Gauss. *L'Enseignement Mathématique*, 30:275–330, 1984.

J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York, N.Y., 1970.

## Appendix A. Code

```
agm<-function(a,b,eps=1e-8,itmax=1000,verbose=TRUE)
{
xold<-max(a,b); yold<-min(a,b); dold<-xold-yold; itel<-1
repeat {
    xnew<-(xold+yold)/2; ynew<-sqrt(xold*yold)
    dnew<-xnew-ynew; rat1<-dnew/dold; rat2<-dnew/(dold^2)
    if (verbose) cat(
        "Iteration: ",formatC(itel,width=3, format="d"),
        "old: ",formatC(c(xold,yold,dold),digits=8,
                 width=12,format="f"),
        "old: ",formatC(c(xnew,ynew,dnew),digits=8,
                 width=12,format="f"),
        "rat: ",formatC(c(rat1,rat2),digits=8,
                 width=12,format="f"),
        "\n")
    if ((dnew < eps) || (itel == itmax))
        return(c(xnew,ynew))
    xold<-xnew; yold<-ynew; dold<-dnew; itel<-itel+1
    }
}
```

Department of Statistics, University of California, Los Angeles, CA 90095-1554

*E-mail address*, Jan de Leeuw: deleeuw@stat.ucla.edu

*URL*, Jan de Leeuw: http://gifi.stat.ucla.edu