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 $D\Gamma(x)$. If $\|D\Gamma(x)\| = \lambda < 1$ we have linear convergence with rate $\lambda$, and if $\|D\Gamma(x)\| = 0$ we have super-linear convergence [Ortega and Rheinboldt, 1970, Chapter 10]. $\|D\Gamma(x)\| = 1$ often indicates sub-linear convergence. Our elementary example below, however, has $\|D\Gamma(x)\| = 1$ and quadratic convergence.

2. THE ARITHMETIC-GEOMETRIC MEAN

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

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

Proof. $0 \leq (\sqrt{a} - \sqrt{b})^2 = 2(AM(a, b) - GM(a, b))$. \qed
From now on suppose, without loss of generality, that \( a > b \). Let \( a_0 = a \) and \( b_0 = b \) and define the sequences

\[
\begin{align*}
(1a) \quad a_n &= \operatorname{AM}(a_{n-1}, b_{n-1}), \\
(1b) \quad b_n &= \operatorname{GM}(a_{n-1}, b_{n-1}).
\end{align*}
\]

**Result 2.** \( a_n > b_n \)

**Proof.** From Result 1. \( \square \)

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

**Result 4.** \( a_\infty = b_\infty \).

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

\[
\begin{align*}
a_\infty &= \operatorname{AM}(a_\infty, b_\infty), \\
b_\infty &= \operatorname{GM}(a_\infty, b_\infty).
\end{align*}
\]

Both equations imply \( a_\infty = b_\infty \). \( \square \)

The common limit \( a_\infty = b_\infty \) is called the arithmetic-geometric mean of \( a \) and \( b \), written as \( \operatorname{AGM}(a, b) \). 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.** \( b < \operatorname{GM}(a, b) < \operatorname{AGM}(a, b) < \operatorname{AM}(a, b) < a \)

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

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

**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 $n$. \[\square\]

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

**Proof.**

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

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

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

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)

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

In fact, we can be even more precise.

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

$$\lim_{n \to \infty} \frac{\delta_n}{\delta_{n-1}^2} = \frac{1}{8} \frac{1}{\text{AGM}(a, b)}.$$
Proof. From Equations (2)
\[
\frac{\delta_n}{\delta_{n-1}^2} = \frac{1}{2} \frac{1}{(\sqrt{a_{n-1}} + \sqrt{b_{n-1}})^2} \rightarrow \frac{1}{8} \text{AGM}(a, b).
\]
\[\square\]

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

**Result 10.** \((a_n - a_{n-1}) \sim -(b_n - b_{n-1}), \text{ i.e.} \)
\[
\lim_{n \to \infty} \frac{a_n - a_{n-1}}{b_n - b_{n-1}} = -1.
\]

Proof.
\[
a_n - a_{n-1} = -\frac{1}{2} (\sqrt{a_{n-1}} - \sqrt{b_{n-1}})(\sqrt{a_{n-1}} + \sqrt{b_{n-1}}),
b_n - b_{n-1} = \sqrt{b_{n-1}}(\sqrt{a_{n-1}} - \sqrt{b_{n-1}}).
\]
and thus
\[
\frac{a_n - a_{n-1}}{b_n - b_{n-1}} = -\frac{1}{2} \frac{\sqrt{a_{n-1}} + \sqrt{b_{n-1}}}{\sqrt{b_{n-1}}} \rightarrow -1.
\]
\[\square\]

3. **Counterexample**

Equation (1) defines a mapping \(\Gamma : \mathbb{R}^2 \rightarrow \mathbb{R}^2\). The derivative of this mapping is
\[
D\Gamma(a, b) = \begin{bmatrix}
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix},
\]
and thus
\[
D\Gamma(\text{AGM}(a, b), \text{AGM}(a, b)) = \begin{bmatrix}
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix},
\]
which has eigenvalues one and zero.

The fact that \(\|D\Gamma(\text{AGM}(a, b), \text{AGM}(a, b))\|_{\infty} = 1\) seems to suggest sub-linear 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


Appendix A. Code

```r
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