Notes on non-convex Lions-Mercier iterations

Jonathan M. Borwein*

January 19, 2009

1 Introduction

Let $R_A(x) := 2P_A(x) - x, R_B(x) := 2P_B(x) - x$, where $P_A, P_B$ denote the Euclidean metric projections, or nearest point maps, on closed sets $A$ and $B$. In our setting, the Lions-Mercier (LM) iteration (which can be given many other names [2] such as Douglas-Rachford or Feinup’s algorithm) is the procedure: reflect, reflect and average:

$$x \mapsto T(x) := \frac{x + R_A(R_B(x))}{2}. \quad (1)$$

Note that a fixed point $z$ of $T$ produces precisely a point $w$ such that $w := P_B(z) = P_A(R_B(z))$ is an element of $A \cap B$. Moreover, if one shows that $\|T(z_n) - z_n\| \to 0$ (known as asymptotic regularity of $z_{n+1} := T(z_n)$) then every cluster point of the corresponding orbit produces a fixed point $z$.

The consequent theory of this and related iterations is well understood in the convex case [1, 2, 3]. In the non-convex case the iteration, also called “divide-and-concur” [5], has been very successful in a variety of reconstruction problems [4, 5] but the theory to explain why is largely absent.

In this note we look at a simple but illustrative special case. The subtlety of this prototype indicates a good deal about the behaviour of the general iteration. Since (LM) has performed much better than other projection iterations on a variety of hard problems [4, 5] we focus on its behaviour.

---

*School of Mathematical and Physical Sciences, University of Newcastle, NSW, 2308, Australia and Faculty of Computer Science, Dalhousie University, Halifax, NS, B3H 2W5, Canada, jborwein@cs.dal.ca, jonathan.borwein@newcastle.edu.au. Supported in part by NSERC and the Canada Research Chair Programme.

1In optical aberration correction as required on the Hubble telescope, however, cyclic projection and its variants have worked well.
2 Dynamics with the circle

In the simplest non-convex case where $B$ is the unit circle and $A$ is a horizontal line of height $\alpha$ the recursion becomes $x_0 := x$, $y_0 := y$ and

\begin{align*}
    x_{n+1} &:= \frac{x_n}{r_n} = \cos \theta_n, \\
y_{n+1} &:= = y_n + \alpha - \frac{y_n}{r_n} = y_n + \alpha - \sin \theta_n,
\end{align*}

where $\theta_n := \arctan(y_n/x_n)$ and $r_n = \sqrt{x_n^2 + y_n^2}$.

Figure 1 shows two steps of the underlying geometric construction. All figures were constructed in Cinderella (www.cinderella.de). A web applet version of the underlying Cinderella construction is available at http://kortenkamps.net/material/IterationBorwein.html. Indeed, many of the insights for the proofs below came from examining the constructions (the number of iterations $N$, the height of the line, and the the initial point are all dynamic—changing one changes the entire visible trajectory).

Let $z_n := (x_n, y_n)$. By symmetry we restrict to $\alpha \geq 0$. It is easy to see that if $x_0 = 0$ then the iteration remains on the vertical axis. We leave this case for the next section.

Thus, we assume that $x_0 > 0$ and it follows from equation (2) that we have $0 < x_n < 1$ for all $n \geq 1$.

We distinguish four cases:

1. $\alpha = 0$. In this case we prove in Theorem 2 below that $z_n \to (1, 0)$. (See Figure 2.)

2. $0 < \alpha < 1$. In this case we conjecture that always $z_n \to (\sqrt{1 - \alpha^2}, \alpha)$. (See Figure 3.)
3. $\alpha = 1$. In this case we prove in Theorem 3 below that $z_n \to (0, y)$ for some finite $y > 1$. (See Figure 4.)

4. $\alpha > 1$. In this infeasible case we prove in Theorem 1 below that $y_n \to \infty$ at linear rate and $x_n \to 0$.

**Theorem 1 (Infeasible case)** If $\alpha > 1$ then $y_n \to \infty$ at linear rate as $n \to \infty$, and $x_n \to 0$.

**Proof.** An easy estimate from equation (3) is $y_{n+1} - y_n \geq \alpha - 1 > 0$. The assertion about $y_n$ follows and the behaviour of $x_n$ is left as an exercise. ■

For the remaining feasible cases the following preliminary computation is useful. We write

$$r_{n+1}^2 = \frac{x_n^2}{r_n^2} + \frac{y_n^2}{r_n^2} + (y_n + \alpha)^2 - 2 \frac{y_n}{r_n} (y_n + \alpha)$$

$$= 1 + \alpha^2 + y_n^2 \left(1 - \frac{2}{r_n}\right) + 2\alpha y_n \left(1 - \frac{1}{r_n}\right).$$

Thus,

$$r_{n+1}^2 - 1 = \alpha^2 + y_n^2 \left(1 - \frac{2}{r_n}\right) + 2\alpha y_n \left(1 - \frac{1}{r_n}\right). \quad (4)$$

**Proposition 1** Suppose that $\alpha = 0$. Suppose also that that $n > 0$ and $r_n > 1$. Then $r_{n+1} < r_n$. 

3
Figure 3: Case with $\alpha = 0.9$.

Figure 4: Case with $\alpha = 1$. 
Proof. Equation (4) becomes

\[ r_{n+1}^2 - 1 = \frac{y_n^2}{r_n^2} (r_n^2 - 2r_n + 1) - \frac{y_n^2}{r_n^2} (r_n - 1)^2. \]

Hence \( r_{n+1}^2 - 1 \leq (r_n - 1)^2 \). Thus, either \( r_{n+1} < 1 \) or \( 0 < r_{n+1} - 1 \leq r_n - 1 \). In either case we are done.

Proposition 2 Suppose that \( \alpha = 0 \) and that \( n > 0 \) and \( r_n < 1 \). Then \( r_{n+1} < 1 \).

Proof. This time we use Equation (4) in the form

\[ 1 - r_{n+1}^2 = \frac{y_n^2}{r_n^2} (2 - r_n) > 0, \]

since \( r_n < 1 \).

Proposition 3 Suppose that \( \alpha = 1 \). Suppose also that that \( n > 0 \) and \( r_n > 2 \). Then \( r_{n+1} < r_n \).

Proof. Equation (4) rewrites as

\[ r_{n+1}^2 - 1 = 1 + \frac{y_n^2}{r_n^2} (r_n^2 - 2r_n) + 2 \frac{y_n}{r_n} (r_n - 1). \]

Hence \( r_{n+1}^2 - 1 < 1 + (r_n^2 - 2r_n) + 2 (r_n - 1) = r_n^2 - 1 \), and \( r_{n+1} < r_n \) as required.

Theorem 2 (Equatorial case) If \( \alpha = 0 \) then \( z_n \to (1, 0) \).

Proof. By Proposition 1 either (a) \( r_n \) strictly decreases to \( r \geq 1 \), which is easily seen to be impossible, or (b) in finitely many steps \( r_n < 1 \). We appeal to Proposition 2 to conclude that \( r_m < 1 \) for all \( m < n \).

We note that

\[ |\tan(\theta_{n+1})| = |1 - r_n| |\tan(\theta_n)| < |\tan(\theta)|, \tag{5} \]

and so \( \tan(\theta_n) \) is decreasing in modulus. It follows, on taking limits in formula (5) that (a) \( r_n \to 1 \) or (b) \( \theta_n \to 0 \). In case (a) we see from equation (3) that \( y_n \to 0 \) and from (2) that \( x_n \to 1 \).

Thus, we are left only with the case that \( \theta_n \to 0 \). But now \( x_{n+1} = \cos(theta_n) \to 1 \) and \( asy_{n+1}/y_{n+1} \to 0 \), the proof is complete.

Theorem 3 (Tangent case) If \( \alpha = 1 \) then \( z_n \to \overline{z} := (0, \overline{y}) \) for some finite \( \overline{y} > 1 \) (and the projection on the sphere of \( \overline{z} \) is the intersection point of the two sets).
**Proof.** An easy estimate from equation (3) is \( y_{n+1} - y_n \geq 0 \). Thence \( y_n \) is nondecreasing with possibly infinite limit \( \overline{y} \). If \( \overline{y} \) is finite then taking limits in (3) shows \( \lim_{n \to \infty} r_n = \lim_{n \to \infty} y_n \), which completes the proof—as \( r \leq 1 \) is easy to rule out.

In the remaining case, by relabeling, we may assume that \( r_n > y_n > 2 \) for all \( n \). Thence Proposition 3 inductively shows that \( r_n \) decreases to some finite \( r > \overline{y} \). This contradiction concludes the proof. ■

**Remark 1** (\( 0 < \alpha < 1 \)) It remains to consider \( 0 < \alpha < 1 \) and it seems probable that similar but more careful arguments using Equation (4) are key to proving the ubiquitous behaviour shown in Figure 3. Noting that \( x > 0 \) is preserved, we observe that it is easy to derive from the iteration formulas (2,3) that for \( n \geq 1 \) we have

\[
y_n - \alpha = (r_{n-1} - 1) \sqrt{1 - x_n^2}.
\] (6)

Hence, \( y_n - \alpha \) and \( r_{n-1} - 1 \) have the same sign, while \( |y_n - \alpha| \leq |r_{n-1} - 1| \). We note that if \( \{x_n\} \) stays bounded away from 1 then \( \varepsilon |r_{n-1} - 1| \leq |y_n - \alpha| \leq (1 - \varepsilon) |r_{n-1} - 1| \) for some \( \varepsilon > 0 \) independent of \( n \).

We also record that

\[
r_{n+1}^2 - 1 = \frac{(y_n + \alpha)^2 (r_n - 1) + (\alpha^2 - y_n^2)}{r_n}.
\] (7)

This shows that if \( y_n - \alpha \) and \( r_n - 1 \) have the opposite sign then \( y_{n+1} - \alpha \) and \( r_{n+1} - 1 \) have the same sign since \( y_{n+1} - \alpha = (1 - 1/r_n) y_n \). In consequence in each, sofar-unproven, spiral only twice are \( y_n - \alpha \) and \( r_n - 1 \) of opposite sign. Moreover, while \( r_n > 1 \) the iteration behaves as in the case of the ball and the line—for which \( T \) is strongly non-expansive—and so \( |r_{n+1} - 1| < |r_n - 1| \).

Also, from the previous two equations, it is clear that the full iteration converges to any feasible cluster point.

### 3 Behaviour on the vertical axis

It is clear both geometrically and analytically that the vertical axis is left invariant by the iteration (2,3). Even so, starting with \( x_0 = 0 \) leads to quite varied behaviour. We note that \( P_B(0) \) is the entire unit disk, and so the mapping is intrinsically multivalued at zero.

Again we distinguish four cases:

1. \( \alpha = 0 \). In this case the mapping has period two for \( y \) in \([-1, 1] \). For \( |y| > 1 \), however \( T^{(2)}([0, y]) = [0, y - 2 \text{sign}(y)] \).

2. \( 0 < \alpha < 1 \). In this case, the behaviour of the map is quite subtle and depends on the the starting point and \( \alpha \). It exhibits periodicity of varied orders when both are rational.
3. $\alpha = 1$. In this case $T([0,y]) = [0,y]$ for $y > 0$ and $T([0,y]) = [0,y+2]$ for $y < 0$. Hence after a finite number of iterations the iteration terminates.

4. $\alpha > 1$. In this infeasible case we again see simpler translational behaviour of $T$.

4 Extensions

Several natural extensions to study (graphically and analytically) take $B$ as the sphere in $n$-dimensional space $E$ and consider:

- $A$ as an affine subspace in $E$ of dimension $2 < m < n$;
- $A$ as a polyhedron (or polyhedral cone) with $n = 2$ or $n = 3$.

Remark 2 Note, even in two dimensions, alternating projections, alternating reflections, project-project and average, and reflect-reflect and average will all often converge to (locally nearest) infeasible points even when $A$ is simply the ray $R := \{[x,0] : x \geq -1/2\}$ and $B$ is the circle as before. They can also behave quite ‘chaotically’. (See Figure 5 for a periodic illustration in Maple.) So the affine nature of the convex set seems quite important.
Remark 3 (Nearest point to an ellipse) Consider the ellipse

\[
E := \left\{ (x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}
\]

in standard form. The best approximation \( P_E(u, v) = \left( \frac{a^2 u}{a^2 - t}, \frac{b^2 v}{b^2 - t} \right) \) where \( t \) solves \( \frac{a^2 u^2}{(a^2 - t)^2} + \frac{b^2 v^2}{(b^2 - t)^2} = 1 \). This generalizes neatly to a hyperbola (one solves the general quartic \( x^4 - ux^3 + vx - 1 = 0 \) and \([x, 1/x]\) is the nearest point.)

Remark 4 (Nearest point to the \( p \)-sphere) For \( 0 < p < \infty \), consider the \( p \)-sphere in two dimensions

\[
S_p := \left\{ (x, y) : |x|^p + |y|^p = 1 \right\}.
\]

Let \( z^* := (1 - z^p)^{1/p} \). For \( uv \neq 0 \), the best approximation \( P_{S_p}(u, v) = (\text{sign}(u)z, \text{sign}(v)z^*) \) where either \( z = 0, 1 \) or \( 0 < z < 1 \) solves

\[
z^{p-1}(z - |u|) - z^{p-1}(z^* - |v|) = 0.
\]

[Then one computes the two or three distances and select the point yielding the least value. It is instructive to make a plot, say for \( p = 1/2 \).] This extends to the case where \( uv = 0 \). Note that this also yields the nearest point formula for the \( p \)-ball.

It should be possible to consider local convergence by linearization of \( T \) from Equation (1). This makes it important to understand approximate solution of a point in the intersection of two hyperplanes.

For the hyperplane \( H_a := \{x : \langle a, x \rangle = b\} \) the projection is

\[
x \mapsto x - \langle a, x \rangle b \|a\|^2.
\]

The consequent averaged-reflection version of the Douglas-Rachford or Lions-Mercier recursion for a point in the intersection of \( N \) distinct hyperplanes is:

\[
x \mapsto x - \frac{2}{N} \sum_{k=1}^{N} \{ \langle a_k, x \rangle - b_k \} \frac{a_k}{\|a_k\|^2}.
\] (8)

The corresponding-averaged projection algorithm is:

\[
x \mapsto x - \frac{1}{N} \sum_{k=1}^{N} \{ \langle a_k, x \rangle - b_k \} \frac{a_k}{\|a_k\|^2}
\] (9)

References


