Effects of a Scaling Factor on the Error Vector in Dykstra's Algorithm

Danica Dralus
Advisor: Professor Don Solmon
Oregon State University REU

August 15, 2001

Abstract

Attempts to calculate the best approximation to a point in a convex set have motivated the development of many algorithms. Dykstra's algorithm converges to the best approximation, but its rate of convergence is not easily determined. This paper focuses on a modified version of Dykstra's algorithm and deals with issues of convergence.

1 Introduction

In physical contexts it is often advantageous to calculate the best approximation of a point in some convex body. Many algorithms utilizing iterative projections on subspaces have been developed to calculate these approximations, each offering its own disadvantages. For example, Bregman's algorithm converges weakly to a point in the intersection of the subspaces but does not necessarily converge to the best approximation, while Dykstra's algorithm always finds the best approximation but does not always offer useful bounds on the rate of convergence. Modifications of such algorithms will hopefully demonstrate the effect of a multiplicative parameter on the convergence of the algorithm.

2 Review

2.1 Spaces

An inner product space is a linear space, denoted throughout this paper as X, in which, for each x, y, $z \in X$ and $\alpha \in \mathbb{R}$, there exists a defined, real scalar $\langle x, y \rangle$ having the following properties:

1. $\langle x, x \rangle \geq 0$

2. $\langle x, x \rangle = 0 \Leftrightarrow x = 0$
3. $\langle x, y \rangle = \langle y, x \rangle$
4. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
5. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

A Hilbert space is a type of inner product space with an additional property.

Definition 1 *A Hilbert space is an inner product space in which every Cauchy sequence converges to a point in the space.*

The norm of x, denoted $||x||$, is defined as

$$||x|| := \sqrt{\langle x, x \rangle}.$$

Many useful estimations involving the norm are known. We will make liberal use of the Schwarz Inequality

$$|\langle x, y \rangle| \leq ||x|| \, ||y||$$

and the triangle inequality

$$||x + y|| \leq ||x|| + ||y||.$$

During discussions of convergence it will be helpful to recall the definition of weak convergence.

Definition 2 *The sequence $\{x_n\}$ converges weakly to x, denoted $x_n \xrightarrow{w} x$, if*

$$\langle x_n, y \rangle \longrightarrow \langle x, y \rangle \qquad \textit{for every} \quad y \in X.$$

Strong convergence implies weak convergence, and they are equivalent when

$$x_n \xrightarrow{w} x \qquad \text{and} \qquad ||x_n|| \longrightarrow ||x||,$$

or when working in a finite dimentional Hilbert space.
The following theorem arises in many proofs.

Theorem 3 *Every bounded sequence in a Hilbert space has a weakly convergent subsequence.*

2.2 Bodies and Projections

The projections of interest are on convex bodies, which are defined below.

Definition 4 *Let K be a subset of X, $x, y \in K$, and $0 \leq \beta \leq 1$. Then K is convex if $\beta x + (1 - \beta) y \in K$.*

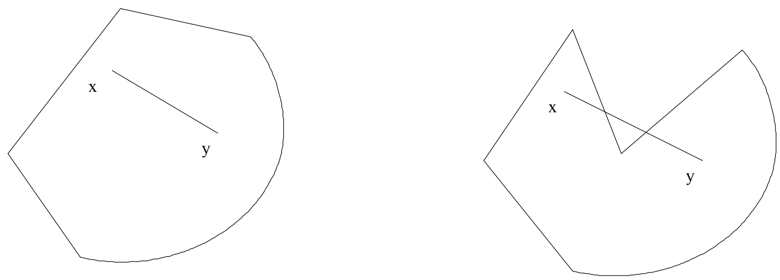

Figure 1: Examples of convex and nonconvex bodies.

From a geometric standpoint, a body is convex if it contains the line segment that joins any two points of the body. Examples of convex and nonconvex bodies are given in Figure 1.

Definition 5 *Let K be a nonempty subset of X and $x \in X$. Then the element $y_0 \in K$ is the best approximation to x from K if*

$$||x - y_0|| = d(x, K),$$

where

$$d(x, K) := \inf ||x - y||.$$

The best approximation to each of two points in a convex set is shown in Figure 2.

The set of all best approximations from x to K is represented by $P_K(x)$, so we have

$$P_K(x) := \{y \in K| \; ||x - y|| = d(x, K)\}.$$

2.3 Bregman's Algorithm

Bregman's algorithm is an iterative process for finding a point in a convex set K by projecting a point $x \in X$ onto each K_i where K is defined as the intersection of a finite number r of closed convex sets $\cap_i^r K_i$. It has been proved that Bregman's algorithm converges weakly to a point in the intersection of the K_i's, but it does not necessarily converge to the best approximation. This follows from the fact that the order of projections is important.

An example of the importance of the order of projections in Bregman's algorithm can be seen in Figure 3. Notice that by projecting onto the x-axis first, the projections converge to the best approximation to x, marked $P1$. If the first projection is onto the unit disc, the algorithm converges to a different point ($P2$) in the intersection of sets that is not the best approximation to x.

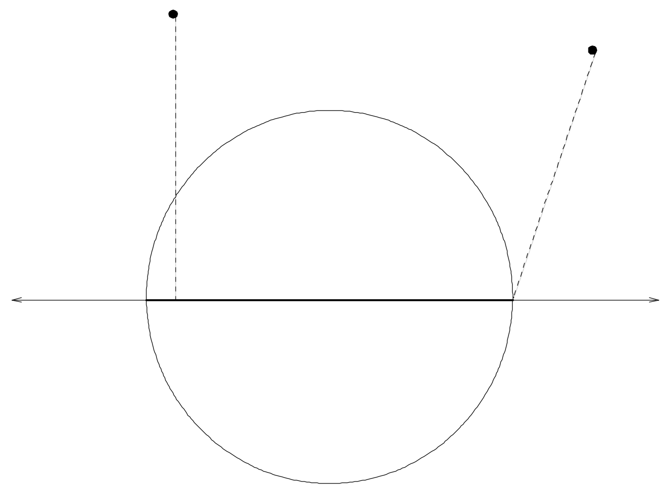

Figure 2: The best approximations to two points in the intersection of the unit disc and the x-axis.

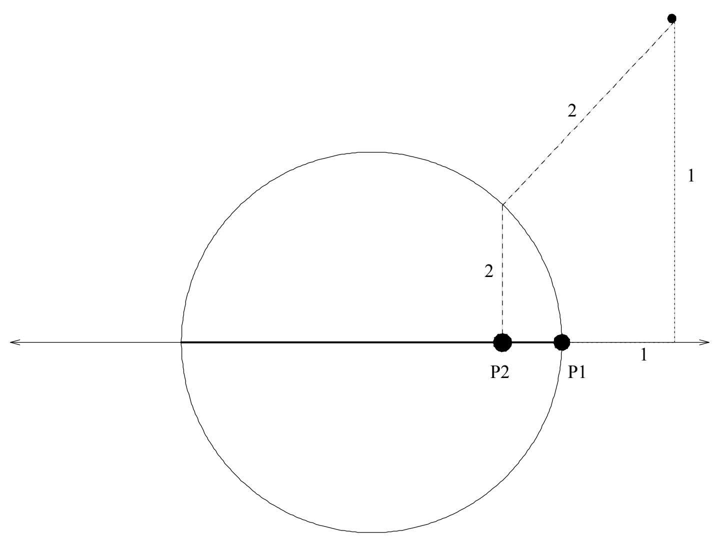

Figure 3: Changing the order of projections in Bregman's algorithm can change the point to which it converges.

2.4 Dykstra's Algorithm

Dykstra's algorithm is an iterative process for finding the best approximation in a convex set K of some $x \in X$ where K is the intersection of a finite number r of closed convex sets $\cap_i^r K_i$. The alogrithm is defined as follows: Let $n \in \mathbb{N}$ with $[n]$ denoting "n mod r", let $P_{K_{[n]}}(x)$ be the projection of x on the subset $K_{[n]}$, and define the recursive relations

$$\begin{aligned} x_0 &:= x \\ e_{-(r-1)} &= \cdots = e_{-1} = e_0 = 0 \\ x_n &:= P_{K_{[n]}}(x_{n-1} + e_{n-r}) \\ e_n &:= x_{n-1} + e_{n-r} - x_n \\ &= x_{n-1} + e_{n-r} - P_{K_{[n]}}(x_{n-1} + e_{n-r}). \end{aligned}$$

It has been proved that Dykstra's algorithm converges strongly to the best approximation of x in K. Bounds on the rate of converges have been explored; they are, however, often difficult to calculate explicitly and frequently so large as to be of little use in applications. In this paper we modify Dykstra's algorithm to see how scaling the error vector e_n affects its convergence.

3 Modified Dykstra Algorithm

Introduce a parameter λ_n to scale the error vector e_n. We will define the algorithm as follows:

$$\begin{aligned} x_0 &:= x \qquad ; \qquad \lambda_n > 0 \\ \lambda_{-(r-1)} &= \cdots = \lambda_{-1} = \lambda_0 = \cdots = \lambda_r = 1 \\ e_{-(r-1)} &= \cdots = e_{-1} = e_0 = 0 \\ x_n &:= P_{K_{[n]}}(x_{n-1} + \lambda_{n-r} e_{n-r}) \\ e_n &:= x_{n-1} + \lambda_{n-r} e_{n-r} - x_n \\ &= x_{n-1} + \lambda_{n-r} e_{n-r} - P_{K_{[n]}}(x_{n-1} + \lambda_{n-r} e_{n-r}). \end{aligned}$$

Note that, when $\lambda_n = 0$ for all possible $n \in \mathbb{N}$, we have Bregman's algorithm. The lemmas and theorems that follow attempt to determine what properties can be attributed to the more general form of Dykstra's algorithm where λ_n is an arbitrary coefficient. At this point in the paper, no assumptions are being made about the magnitude of λ_n or its dependencies, except that it does not depend on the point x. As properties of the modified Dykstra algorithm (MoDA) are explored, limitations on the λ_n's will be introduced. Lemmas and theorems will correspond to theorems in Chapter 9 of Deutsch's book *Best Approximation in Inner Product Spaces* [2].

Lemma 6 *For each* $n \in \mathbb{N}$

$$\langle x_n - y,\, \lambda_n e_n \rangle = \lambda_n \langle x_n - y,\, e_n \rangle \geq 0 \quad \textit{for all} \quad y \in K_{[n]}.$$

Proof. It can be shown easily that

$$\langle x - y_0,\, y - y_0 \rangle \leq 0 \text{ for all } y \in K$$

(See Deutsch, Theorem 4.1). Using property four of an inner product space, we have

$$\langle y_0 - y,\, x - y_0 \rangle \geq 0 \text{ for all } y \in K.$$

Apply the aforementioned theorem with

$$x = x_{n-1} + \lambda_{n-r} e_{n-r}, \qquad y_0 = P_{[n]}(x_{n-1} + \lambda_{n-r} e_{n-r}),$$

and $K = K_{[n]}$. Then we have

$$\begin{aligned}\langle x_n - y,\, \lambda_n e_n \rangle &= \langle P_{[n]}(x_{n-1} + \lambda_{n-r} e_{n-r}) - y,\, \lambda_n \left[x_{n-1} + \lambda_{n-r} e_{n-r} + P_{[n]}(x_{n-1} + \lambda_{n-r} e_{n-r}) \right] \rangle \\ &= \lambda_n \langle y_0 - y,\, x - y_0 \rangle.\end{aligned}$$

Thus $\langle x_n - y,\, \lambda_n e_n \rangle = \lambda_n \langle x_n - y,\, e_n \rangle \geq 0$ for all $y \in K_{[n]}$. ■

Lemma 7 *For each $n \geq 0$,*

$$\begin{aligned} x - x_n &= \sum_{i=0}^{n} \left(e_i - \lambda_{i-r} e_{i-r} \right), \text{ alternately written} \\ &= \sum_{i=n-(r-1)}^{n} e_i + \sum_{i=0}^{n-r} (1 - \lambda_i) e_i. \end{aligned}$$

Proof. By induction on n. For $n = 0$, $x - x_0 = x - x = 0$, and since $e_n = 0$ for $n \leq 0$, the sum is zero as well. Assume that the result is valid for some $n \geq 0$. Then

$$\begin{aligned} x - x_{n+1} &= (x - x_n) + (x_n - x_{n+1}) \\ &= \sum_{i=0}^{n} \left(e_i - \lambda_{i-r} e_{i-r} \right) + e_{n+1} - \lambda_{n+1-r} e_{n+1-r} \\ &= \sum_{i=0}^{n+1} \left(e_i - \lambda_{i-r} e_{i-r} \right). \end{aligned}$$

This shows that the assumption is valid for $n + 1$. To write this in a more useful form, perform a change of indices.

$$\begin{aligned}\sum_{i=0}^{n}(e_i - \lambda_{i-r}e_{i-r}) &= \sum_{i=0}^{n} e_i - \sum_{i=0}^{n}\lambda_{i-r}e_{i-r}\\ &= \sum_{i=0}^{n} e_i - \sum_{j=-r}^{n-r}\lambda_j e_j\\ &= \sum_{i=0}^{n} e_i - \sum_{j=0}^{n-r}\lambda_j e_j\\ &= \sum_{i=0}^{n-r} e_i - \sum_{j=0}^{n-r}\lambda_j e_j + \sum_{i=n-(r-1)}^{n} e_i\\ &= \sum_{i=n-(r-1)}^{n} e_i + \sum_{i=0}^{n-r}(1-\lambda_i)e_i.\end{aligned}$$

■

Lemma 8 *For each $n \in \mathbb{N}$, $0 \le m \le n$, and $y \in K$,*

$$||x_m - y||^2 = ||x_n - y||^2 + \sum_{k=m+1}^{n} ||x_k - x_{k-1}||^2 + 2\sum_{k=m+1}^{n}\langle \lambda_{k-r}e_{k-r},\, x_{k-r} - x_k\rangle + 2\sum_{k=n-(r-1)}^{n}\langle e_k,\, x_k - y\rangle - 2\sum_{k=m-(r-1)}^{n}\langle \lambda_k e_k,\, x_k - y\rangle.$$

Proof. Following Deutsch's proof of Lemma 9.19, we see

$$||y_m - y_{n+1}||^2 = ||y_n - y_{n+1}||^2 + \sum_{k=m+1}^{n}||y_{k-1} - y_k||^2 + 2\sum_{i=m+1}^{n}\left(\sum_{j=i+1}^{n+1}\langle y_{i-1} - y_i,\, y_{j-1} - y_j\rangle\right).$$

It still follows that

$$\sum_{i=m+1}^{n}\left(\sum_{j=i+1}^{n+1}\langle y_{i-1} - y_i,\, y_{j-1} - y_j\rangle\right) = \sum_{i=m+1}^{n}\langle y_{i-1} - y_i,\, y_i - y_{n+1}\rangle.$$

We can substitute $y_i = x_i$ for all $i \le n$ and $y_{n+1} = y$ into the above equation to find

$$\begin{aligned}
&\sum_{i=m+1}^{n} \langle y_{i-1} - y_i,\, y_i - y_{n+1}\rangle \\
&\quad = \sum_{i=m+1}^{n} \langle x_{i-1} - x_i,\, x_i - y\rangle \\
&\quad = \sum_{i=m+1}^{n} \langle e_i - \lambda_{i-r} e_{i-r},\, x_i - y\rangle \\
&\quad = \sum_{i=m+1}^{n} \langle e_i,\, x_i - y\rangle - \sum_{i=m+1}^{n} \langle \lambda_{i-r} e_{i-r}\,, x_i - y\rangle \\
&\quad = \sum_{i=m+1}^{n} \langle e_i,\, x_i - y\rangle - \sum_{i=m+1}^{n} [\langle \lambda_{i-r} e_{i-r}\,, x_i - x_{i-r}\rangle + \langle \lambda_{i-r} e_{i-r}\,, x_{i-r} - y\rangle] \\
&\quad = \sum_{i=m+1}^{n} \langle e_i,\, x_i - y\rangle - \sum_{i=m+1}^{n} \langle \lambda_{i-r} e_{i-r}\,, x_{i-r} - y\rangle + \sum_{i=m+1}^{n} \langle \lambda_{i-r} e_{i-r}\,, x_{i-r} - x_i\rangle \\
&\quad = \sum_{i=m+1}^{n} \langle e_i,\, x_i - y\rangle - \sum_{i=m+1-r}^{n-r} \langle \lambda_i e_i, x_{i-r} - y\rangle + \sum_{i=m+1}^{n} \langle \lambda_{i-r} e_{i-r}\,, x_{i-r} - x_i\rangle \\
&\quad = \sum_{i=n-r+1}^{n} \langle e_i,\, x_i - y\rangle - \sum_{i=m-r+1}^{m} \langle \lambda_i e_i\,, x_i - y\rangle + \sum_{i=m+1}^{n} \langle \lambda_{i-r} e_{i-r}\,, x_{i-r} - x_i\rangle.
\end{aligned}$$

Further substitution yields

$$\begin{aligned}
||x_m - y||^2 = ||x_n - y||^2 + \sum_{k=m+1}^{n} ||x_{k-1} - x_k||^2 + 2 \sum_{k=n-r+1}^{n} \langle e_k,\, x_k - y\rangle \\
- 2 \sum_{k=m-(r-1)}^{n} \langle \lambda_k e_k,\, x_k - y\rangle + \sum_{k=m+1}^{n} \langle \lambda_{k-r} e_{k-r},\, x_{k-r} - x_k\rangle,
\end{aligned}$$

which completes the proof. ■

Lemma 9 $\{x_n\}$ *is a bounded sequence and*

$$\sum_{1}^{\infty} ||x_{k-1} - x_k||^2 < \infty.$$

In particular,

$$||x_0 - y||^2 \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Proof. Noting that the third and fourth terms of Lemma 8 are non-negative by Lemma 6, this proof follows directly from Deutsch's Lemma (9.20) and has therefore been omitted. ■

Lemma 10 *For any $n \in \mathbb{N}$ and $0 < \lambda_n \leq 1$,*

$$||e_n|| \leq ||x_{n-1} - x_n|| + |\lambda_{n-r}| \sum_{k=2}^{n} ||x_{k-2} - x_{k-1}|| \leq \sum_{k=1}^{n} ||x_{k-1} - x_k||.$$

Proof. By induction on n. For $n = 1$, we have

$$\begin{aligned} ||e_1|| &= ||x_0 - x_1 + \lambda_{1-r} e_{1-r}|| = ||x_0 - x_1|| \\ &= ||x_{1-1} - x_1|| + |\lambda_{1-1}| \sum_{k=2}^{1} ||x_{k-2} - x_{k-1}|| \\ &= \sum_{k=1}^{1} ||x_{k-1} - x_k||, \end{aligned}$$

since $e_k = 0$ for all $k \leq 0$. Assume that the relation is true for all $m \leq n$. Then we have

$$\begin{aligned} ||e_{n+1}|| &= ||x_n - x_{n+1} + \lambda_{n+1-r} e_{n+1-r}|| \\ &\leq ||x_n - x_{n+1}|| + ||\lambda_{n+1-r} e_{n+1-r}|| \\ &\leq ||x_n - x_{n+1}|| + |\lambda_{n+1-r}| \, ||e_{n+1-r}|| \\ &\leq ||x_n - x_{n+1}|| + |\lambda_{n+1-r}| \left(||x_{n+1-r-1} - x_{n+1-r}|| + |\lambda_{n+1-r-r}| \sum_{k=2}^{n+1-r} ||x_{k-2} - x_{k-1}|| \right) \\ &\leq ||x_n - x_{n+1}|| + |\lambda_{n+1-r}| \left(||x_{n-r} - x_{n+1-r}|| + \sum_{k=2}^{n+1-r} ||x_{k-2} - x_{k-1}|| \right) \\ &\leq ||x_n - x_{n+1}|| + |\lambda_{n+1-r}| \sum_{k=2}^{n+1} ||x_{k-2} - x_{k-1}|| \\ &\leq ||x_n - x_{n+1}|| + \sum_{k=2}^{n+1} ||x_{k-2} - x_{k-1}|| \\ &= \sum_{k=1}^{n+1} ||x_{k-1} - x_k||. \end{aligned}$$

■

Lemma 11 *For $0 < \lambda_n \leq 1$,*

$$\liminf_{n} \sum_{k=n-(r-1)}^{n} |\langle x_k - x_n, \, e_k \rangle| = 0.$$

Proof. This proof follows directly from Deutsch and has therefore been omitted. ■

At this point in our string of analogues, we find a lemma that does not follow exactly. For Dykstra's algorithm, we can show that there exists a subsequence x_{n_j} of x_n such that

$$\limsup_j \langle y - x_{n_j},\, x - x_{n_j}\rangle \leq 0 \text{ for each } y \in K, \text{ and}$$

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_n,\, e_k\rangle| = 0.$$

For MoDA, however, we have the inequality

$$\langle y - x_n,\, x - x_n\rangle \leq \sum_{i=n-(r-1)}^{n} \langle x_i - x_n,\, e_i\rangle + \sum_{i=0}^{n-r}(1-\lambda_i)\,\langle x_i - x_n,\, e_i\rangle,$$

which contains an infinite sum when $n \longrightarrow \infty$. Thus we consider a different approach.

Theorem 12 *Let* $0 < \lambda < 1$, λ *constant. Then*

$$||\, e_n \,|| \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty.$$

Proof. Let $1 \leq j < r$, $m > 0$, $0 < \lambda < 1$ with λ constant, and let $n = mr+j$. Then we can write the inequalities

$$\begin{aligned} ||\, e_n \,|| &= ||\, e_{mr+j} \,|| = ||\, x_{n-1} - x_n + \lambda e_{n-r} \,|| \\ &\leq ||\, x_{n-1} - x_n \,|| + \lambda ||\, e_{n-r} \,|| \\ &\leq ||\, x_{n-1} - x_n \,|| + \lambda ||\, x_{n-r-1} - x_{n-r} \,|| + \lambda^2 ||\, e_{n-2r} \,|| \\ &= ||\, x_{mr+j-1} - x_{mr+j} \,|| + \lambda ||\, x_{(m-1)r+j-1} - x_{(m-1)r+j} \,|| + \lambda^2 ||\, e_{(m-2)r+j} \,||. \end{aligned}$$

In general, for $1 \leq i < m$, we have

$$\begin{aligned} ||\, e_n \,|| &= ||\, e_{mr+j} \,|| \\ &\leq ||\, x_{mr+j-1} - x_{mr+j} \,|| + \lambda ||\, x_{(m-1)r+j-1} - x_{(m-1)r+j} \,|| + \lambda^2 ||\, e_{(m-2)r+j} \,|| \\ &\quad + \cdots + \lambda^{i-1} ||\, x_{(m-i+1)r+j-1} - x_{(m-i+1)r+j} \,|| + \lambda^i ||\, e_{(m-i)r+j} \,|| \\ &= \sum_{l=0}^{i-1} \left[\lambda^l ||\, x_{(m-l)r+j-1} - x_{(m-l)r+j} \,||\right] + \lambda^i ||\, e_{(m-i)r+j} \,||. \end{aligned}$$

We know that $||\, e_n \,||$ is bounded by Lemma 10, so $\{\, e_n \,\}$ is bounded, meaning there exists some $M > 0$ such that $||\, e_n \,|| \leq M < \infty$ for all n. Let $\varepsilon > 0$. Since $0 < \lambda < 1$, there's $K_1 < 0$ such that

$$\lambda^a < \frac{\varepsilon}{2M} \text{ whenever } a \geq K_1.$$

Since $\sum_{k=1}^{\infty} ||\, x_{k-1} - x_k \,||^2 < \infty$, there exists some $K_2 > 0$ such that

$$\sum_{k=b}^{\infty} ||\, x_{k-1} - x_k \,||^2 < \frac{(1-\lambda^2)\,\varepsilon^2}{4} \text{ whenever } b \geq K_2$$

Suppose that $n = mr + j$ with $m - K_2 > K_1$. Then

$$||\, e_n \,|| \leq \sum_{l=0}^{m-K_2-1} \left[\lambda^l ||\, x_{(m-l)r+j-1} - x_{(m-l)r+j} \,||\right] + \lambda^{m-K_2} ||\, e_{(m-(m-K_2)r+j} \,||.$$

Since $m - K_2 > K_1$,

$$\lambda^{m-K_2} < \frac{\varepsilon}{2M} \text{ and}$$
$$\lambda^{m-K_2} ||\, e_{(m-(m-K_2)r+j} \,|| < \frac{\varepsilon}{2}.$$

The other term is bounded by

$$\begin{aligned}
&\sum_{l=0}^{m-K_2-1} \lambda^l ||\, x_{(m-l)r+j-1} - x_{(m-l)r+j} \,|| \\
&\quad \leq \left[\sum_{l=0}^{m-K_2-1} \lambda^{2l}\right]^{\frac{1}{2}} \left[\sum_{l=0}^{m-K_2-1} ||\, x_{(m-l)r+j-1} - x_{(m-l)r+j} \,||^2\right]^{\frac{1}{2}} \\
&\quad \leq \left[\sum_{l=0}^{\infty} \lambda^{2l}\right]^{\frac{1}{2}} \left[\sum_{l=K_1}^{\infty} ||\, x_{lr+j-1} - x_{lr+j} \,||^2\right]^{\frac{1}{2}} \\
&\quad \leq \left[\frac{1}{1-\lambda^2}\right]^{\frac{1}{2}} \left[\sum_{l=K_1}^{\infty} ||\, x_{l-1} - x_l \,||^2\right]^{\frac{1}{2}} \\
&\quad \leq \frac{1}{\sqrt{1-\lambda^2}} \left[\frac{(1-\lambda^2)\,\varepsilon^2}{4}\right]^{\frac{1}{2}} \\
&\quad < \frac{\varepsilon}{2}.
\end{aligned}$$

Thus, we have

$$||\, e_n \,|| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which can be made arbitrarily small. ■

The previous theorem could also be proved using the following lemma.

Lemma 13 *Let $0 < \lambda < 1$, λ constant. Then*

$$\sum_{n=1}^{\infty} ||e_n||^2 < \infty.$$

Proof. Let $1 \leq j < r$, $m > 0$, $0 < \lambda < 1$ with λ constant, and let $n = mr+j$. It is easy to show that, in general,

$$||e_n|| \leq \sum_{l=0}^{m} \lambda^l ||x_{(m-l)r+j-1} - x_{(m-l)r+j}||.$$

Using this estimation for $||e_n||$ along with the Minkowski inequality for integrals [3, p.4], we can show

$$\left[\sum_{m=0}^{\infty} ||e_n||^2\right]^{\frac{1}{2}} \leq \left[\sum_{m=0}^{\infty}\left(\sum_{l=0}^{m} \lambda^l ||x_{(m-l)r+j-1} - x_{(m-l)r+j}||\right)^2\right]^{\frac{1}{2}}$$

$$\leq \sum_{l=0}^{m}\left[\sum_{m=0}^{\infty} \lambda^{2l} ||x_{(m-l)r+j-1} - x_{(m-l)r+j}||^2\right]^{\frac{1}{2}}$$

$$\leq \left[\sum_{l=0}^{m} \lambda^l\right]\left[\sum_{m=0}^{\infty} ||x_{m+1} - x_m||^2\right]^{\frac{1}{2}}$$

When $0 < \lambda < 1$, the first product converges, and the second product was shown to be bounded by Lemma 9. Thus we have

$$\sum_{n=1}^{\infty} ||e_n||^2 < \infty.$$

■

4 Conclusion

We have not yet been able to prove the convergence of MoDA; however, weak convergence appears highly probable. One argument for this expectation is geometric. Suppose λ is a very small constant, and consider the projection of a point x onto the unit circle and the x-axis (See Figure 4). Notice that, if we project onto the unit circle first, we converge to a point between the Bregman approximation and the Dykstra approximation. We expect to have at least weak convergence since Bregman's algorithm converges weakly. If we project onto the x-axis first, then we converge to the best approximation. Thus the order of projections is important, as in Bregman's algorithm. Whether or not an expression relating λ to the point of convergence of the modified Dykstra algorithm relative to either Bregman's or Dykstra's points has yet to be determined.

5 References

References

[1] L.M. Bregman, *The Method of Successive Projection for Finding a Common Point of Convex Sets*, Sov. Math. Dok. 6, 1965.

[2] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer, New York, 2001.

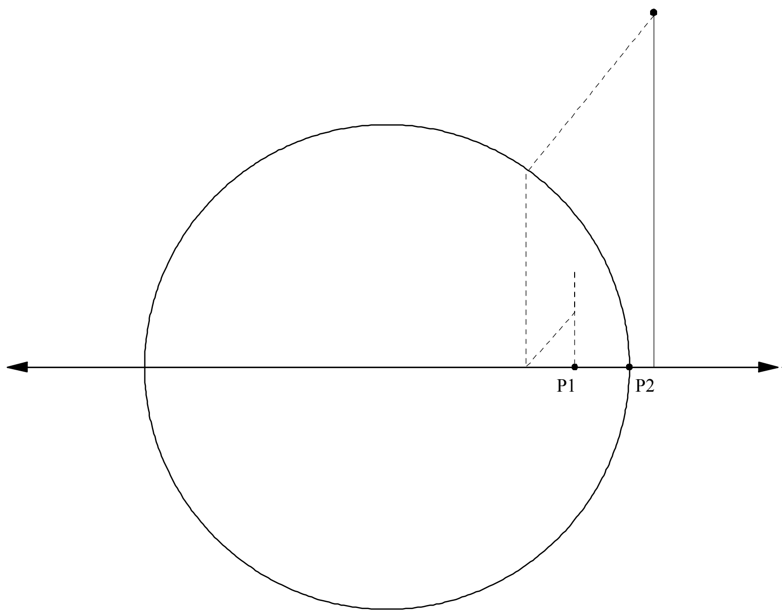

Figure 4: Example of MoDA convergence. Note that order of projections is important.

[3] B. Petersen, *Introduction to the Fourier Transform & Pseudo-differential Operators*, Pitman, Boston, 1983.