MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 117, JANUARY, 1972

Rate of Convergence of Lawson's Algorithm

By A. K. Cline

Abstract. The algorithm of Charles L. Lawson determines uniform approximations of functions as limits of weighted L_2 approximations. Lawson noticed from experimental evidence that the algorithm seemed to converge linearly and convergence was related to a factor which was the ratio of the largest nonmaximum error of the best uniform approximation to the maximum error. This paper proves the linear convergence and explores the relation of the rate of convergence to this ratio.

1. Introduction. In his Ph.D. dissertation of 1961, Charles L. Lawson discussed an algorithm for solving uniform approximation problems by means of limits of weighted p-norm solutions. Since then, this algorithm has been explored further by several authors. The algorithm is mentioned in Rice [3], and a variation on Lawson's algorithm was shown to produce p-norm approximations ($p > 2$) as a limit of weighted L_2 norm solutions in Rice and Usow [4]. In the Ph.D. dissertation of this author [1], Lawson's algorithm (originally defined for approximation on finite sets) was extended to the case of approximation on compact Hausdorff spaces. Presently, attempts are underway extending Lawson's algorithm in a different fashion for solving L_1 approximation problems.

In his dissertation, Lawson gave conditions for convergence of the weighted L_p solutions to the uniform solution. In some cases (theoretically possible but computationally highly unlikely), the algorithm may have to be restarted a finite number of times before it converges to the proper solution. When it converges to the uniform solution, Lawson noticed experimental results indicating linear convergence with a convergence factor linked to a certain ratio of error at a point to maximum error of the solution.

It is the purpose of this paper to show that the Lawson algorithm does have linear convergence and demonstrate the importance of the convergence factor which Lawson noticed experimentally. In Section 2, the basic theory of weighted L_2 approximations will be given, the algorithm introduced, and conditions on its convergence to the uniform solution given. In Section 3, the fundamental rate of convergence results are proved through a series of lemmas.

2. Description of the Lawson Algorithm. Although the algorithm was defined for approximation of vector-valued functions by means of weighted L_p approximations, in this paper we consider only real-valued functions and weighted L_2 approximations. We assume we are given a finite set $X = \{x_i\}_{i=1}^m$, a function f on X, and a linear space of approximations G. We let n be the dimension and assume f is not contained in G (hence $n + 1 \leqq m$). We seek to find an element $g^* \in G$ such that

Received May 14, 1971.

AMS 1969 *subject classifications.* Primary 4141; Secondary 4140, 6520.

Key words and phrases. Uniform approximation, least squares approximation.

Copyright © 1972, American Mathematical Society

$$\|g^* - f\| \leqq \|g - f\| \quad \text{for all } g \in G,$$

where $\|\cdot\|$ indicates the uniform norm. We further assume that G has the Chebyshev property (i.e., no element of G has n zeros on X other than the identically zero function), which guarantees that there exists a unique such best uniform approximation g^* to f.

Given a nonnegative, unit weight function w on X (i.e., $\sum_{j=1}^{m} w_j = 1$ and $w_j \geqq 0$ for $j = 1, \cdots, m$), we seek a best weighted L_2 approximation to f as the $\tilde{g} \in G$ such that

$$\left(\sum_{j=1}^{m} w_j[f(x_j) - \tilde{g}(x_j)]^2\right)^{1/2} \leqq \left(\sum_{j=1}^{m} w_j[f(x_j) - g(x_j)]^2\right)^{1/2}$$

for all $g \in G$. A characterization of such solutions is given by the standard orthogonality property:

$\tilde{g}$ minimizes the quantity $(\sum_{j=1}^{m} w_j[f(x_j) - g(x_j)]^2)^{1/2}$, over all $g \in G$ if and only if $\sum_{j=1}^{m} w_j(f(x_j) - \tilde{g}(x_j))g(x_j) = 0$ for all $g \in G$.

Using the orthogonality property, it is easy to show that such a unique best w-weighted L_2 approximation $\tilde{g}$ to f exists if and only if there are at least n positive components of the weight function w. We label the set of all such weight functions W. That is,

$$W = \left\{ w = \{w_j\}_{j=1}^{m} : \sum_{j=1}^{m} w_j = 1,\ w_j \geqq 0 \text{ for all } j \text{ and } > 0 \text{ for at least } n \text{ values of } j \right\}.$$

It is clear that W is not a compact set; however, it is the union of a countable collection of compact sets. To see this, for $\epsilon > 0$, let

$$W_\epsilon = \{w = \{w_j\}_{j=1}^{m} : w \in W \text{ and } w_j \geqq \epsilon \text{ for at least } n \text{ values of } j\}.$$

Then, for any sequence $\{\epsilon_i\}_{i=1}^{\infty}$ with limit zero, it is clear that $W = \bigcup_{i=1}^{\infty} W_{\epsilon_i}$. The compactness of such sets W_ϵ will be exploited in Section 3.

To summarize results to this point, we have that for $w \in W$ there exists a unique best w-weighted L_2 approximation $\tilde{g}$ to f. We denote the mapping of w to $\tilde{g}$ by B. That is,

$$B: W \rightarrow G \quad \text{with } \tilde{g} = B(w).$$

We now introduce a mapping $F: W \rightarrow W$. For $w \in W$, determine $\tilde{g} = B(w)$ and let $r = f - \tilde{g}$ be the residual function. Define a new weight $w' = F(w)$ such that, for $j = 1, \cdots, m$,

$$w'_j = w_j\,|r_j| \Big/ \sum_{i=1}^{m} w_i|r_i|,$$

where $r_j = r(x_j)$. That $w' \in W$ is shown in Lawson [2, p. 70]. We are now in a position to define the algorithm.

0. Let $k = 0$ and select $w^0 \in W$.
1. Determine $\tilde{g}^k = B(w^k)$ and let $\sigma^k = (\sum_{j=1}^{m} w_j^k\,|r_j^k|^2)^{1/2}$.

2. If $\sigma^k = \sigma^{k-1}$ then stop; otherwise let $w^{k+1} = F(w^k)$, increase k by 1, and return to step 1.

Lawson showed the sequence $\{\sigma^k\}_{k=0}^{\infty}$ to be increasing and bounded above by $\tau^* = ||f - g^*||$. In the case that $\sigma^k \to \tau^*$ as $k \to \infty$, the sequence $\{\bar{g}^k\}_{k=0}^{\infty}$ of weighted L_2 approximations has limit g^*, the best uniform solution. To guarantee this convergence, it is necessary and sufficient to assume that for some "approximator determining set" (or "critical set") E_0, every weight function w^k is positive on every point of E_0. An approximator determining set is a subset of the extremal set $E = \{x \in X: |f(x) - g^*(x)| = ||f - g^*||\}$, on which g^* is also the best uniform approximation to f. With the assumptions that all functions are real-valued and that G is a Chebyshev system, we are guaranteed the existence of some approximator determining set of exactly $n + 1$ points. E is always an approximator determining set and, if E contains exactly $n + 1$ points, is the only such set.

The assumption that there exists such an E_0 on which every w^k is strictly positive is, in practical consideration, not at all strong, although examples can be produced where this is violated and, hence, $\{\bar{g}^k\}$ does not have limit g^* (see Lawson [2, p. 83]).

Henceforth, it will be assumed that the algorithm does converge to the uniform solution. That is, $\bar{g}_k \to g^*$ and $\sigma^k \to \tau^*$ as $k \to \infty$.

Lawson reported that according to numerical experiments the convergence of $\{\sigma^k\}$ to τ^* was related to the constant

$$\rho = \max\{|f(x) - g^*(x)|: x \notin E\}/\tau^*.$$

In fact he observed that

$$(\tau^* - \sigma^k)/(\tau^* - \sigma^{k-1}) \to \rho, \quad \text{and also} \quad (\tau^k - \tau^*)/(\tau^{k-1} - \tau^*) \to \rho,$$

where $\tau^k = ||f - \bar{g}^k||$.

It will be shown here that the algorithm does converge in a linear fashion and that the factor of convergence is at most ρ. To be specific, for every $\lambda > \rho$ there is an $M > 0$ such that for all k, $||g^* - \bar{g}^k|| \leqq M\lambda^k$ and $\tau^k - \tau^* \leqq M\lambda^k$. This result is given as Theorem 2.

3. Rate of Convergence. This section presents the proof of the convergence result stated above. For ease of understanding, the proof has been split into five lemmas, two theorems, and two corollaries. In order to convey the importance of each subresult, an outline of this section is presented.

First, it will be shown in Lemmas 1 and 2 that the operators B and F satisfy Lipschitz continuity conditions on the compact sets W_ϵ. That B and F are simply continuous on W is not difficult to show, but a stronger result is required and this stronger result does not hold on all of W. For this reason, the compact sets W_ϵ are considered and prove sufficient for later application.

In Lemma 3, it is shown that points not in the extremal set E have weights tending to zero. This is used to prove Lemma 4: that the quantity $\sum_{i=1}^{m} w_i|r_i|$ in the denominator of the definition of F tends to the constant τ^*. These two results are used to show that the rate of convergence of weights at a point x to zero, mentioned in Lemma 3, is in fact linear with convergence factor related to the ratio $|f(x) - g^*(x)|/\tau^*$. This is given as Lemma 5. The maximum of such ratios is exactly the quantity ρ, thus ρ governs the convergence of the total weight of the set $X \sim E$ to zero.

Theorem 1 and its two corollaries show that if all weight is concentrated on the set E and if the residual functions r^k are sufficiently close to the best uniform residual $r^* = f - g^*$, then the algorithm converges immediately. Theorem 2 links all these ideas by taking the sequence $\{w^k\}$ and defining a new weight sequence $\{\bar{w}^k\}$. Elements of the sequence $\{\bar{w}^k\}$ have all weight concentrated on E which may not be the case for elements of $\{w^k\}$, but the two sequences grow closer with increasing k. The degree of closeness is determined by Lemma 5. Then the Lipschitz continuity conditions of Lemmas 1 and 2 are applied to obtain the desired rate of convergence of $\{r^k - r^*\}$.

LEMMA 1. *Let $\bar{w}^1$, $\bar{w}^2 \in W_\epsilon$ where $\epsilon > 0$. Then there exists a constant M_B (depending only upon ϵ) such that*

$$||\bar{g}^1 - \bar{g}^2|| \leqq M_B \, ||\bar{w}^1 - \bar{w}^2|| \quad \textit{where } \bar{g}^1 = B(\bar{w}^1) \textit{ and } \bar{g}^2 = B(\bar{w}^2).$$

Proof. Let D_1 and D_2 be $m \times m$ diagonal matrices with elements $\{\bar{w}_i^1\}_{i=1}^m$ and $\{\bar{w}_i^2\}_{i=1}^m$ respectively. Select a basis $\{g_i\}_{i=1}^m$ for G and let A be the $m \times n$ matrix with elements

$$(A)_{i,j} = g_j(x_i), \qquad i = 1, \cdots, m, j = 1, \cdots, n,$$

and b be the m-vector with elements

$$(b)_i = f(x_i), \qquad i = 1, \cdots, m.$$

Now, expanding the solutions $\bar{g}^1$ and $\bar{g}^2$ in terms of the basis:

$$\bar{g}^1 = \sum_{j=1}^n \alpha_j^1 g_j \quad \text{and} \quad \bar{g}^2 = \sum_{j=1}^n \alpha_j^2 g_j,$$

from the orthogonality property it follows that the m-vectors α^1 and α^2 satisfy $\alpha^1 = (A^T D_1 A)^{-1} A^T D_1 b$ and $\alpha^2 = (A^T D_2 A)^{-1} A^T D_2 b$. Thus,

$$\begin{aligned}\alpha^1 - \alpha^2 &= [(A^T D_1 A)^{-1} - (A^T D_2 A)^{-1}] A^T D_1 b + (A^T D_2 A)^{-1} A^T (D_1 - D_2) b \\ &= (A^T D_2 A)^{-1} [(A^T D_2 A) - (A^T D_1 A)](A^T D_1 A)^{-1} A^T D_1 b \\ &\quad + (A^T D_2 A)^{-1} A^T (D_1 - D_2) b \\ &= (A^T D_2 A)^{-1} A^T (D_2 - D_1)[A(A^T D_1 A)^{-1} A^T D_1 - I] b.\end{aligned}$$

Letting $||\cdot||_1$ indicate the L_1 vector and subordinate matrix norm, we have

$$\begin{aligned}||\alpha^1 - \alpha^2||_1 \leqq\ & ||(A^T D_2 A)^{-1}||_1 \cdot ||A^T||_1 \cdot ||D_2 - D_1||_1 \\ & \cdot (||A||_1 \cdot ||(A^T D_1 A)^{-1}||_1 \cdot ||A^T||_1 \, ||D_1||_1 + 1) \cdot ||b||_1.\end{aligned}$$

Furthermore, $||D_2 - D_1||_1 = ||\bar{w}_1 - \bar{w}_2||$ and $||D_1||_1 \leqq 1$.

Since W_ϵ is a compact set on which $||(A^T D A)^{-1}||_1 < \infty$ for every diagonal matrix D corresponding to a weight $w \in W_\epsilon$, the quantity $||(A^T D A)^{-1}||_1$ is uniformly bounded on W_ϵ.

This implies the existence of a constant M_1 depending only on ϵ such that

$$||\alpha^1 - \alpha^2||_1 \leqq M_1 \, ||\bar{w}^1 - \bar{w}^2||.$$

The proof is completed by letting $M_B = M_1 \cdot \max_j ||g_j||$ and noticing that

$$||\bar{g}^1 - \bar{g}^2|| = \left|\left|\sum_{j=1}^{n} (\alpha_j^1 - \alpha_j^2)\cdot g_j\right|\right| \leqq \sum_{j=1}^{n} |\alpha_j^1 - \alpha_j^2| \, ||g_j||$$
$$\leqq \max_i ||g_i||\cdot||\alpha^1 - \alpha^2||_1 \leqq \max_i ||g_i||\cdot M_1\cdot||\bar{w}^1 - \bar{w}^2||. \quad \square$$

LEMMA 2. *Let $\bar{w}_1, \bar{w}_2 \in W_\epsilon$ where $\epsilon > 0$, then there exists a constant M_F (depending only upon ϵ) such that*

$$||F(\bar{w}^1) - F(\bar{w}^2)|| \leqq M_F \, ||\bar{w}^1 - \bar{w}^2||.$$

Proof. First, notice that, for $\beta_1, \beta_2 \neq 0$,

$$\left|\frac{\alpha_1}{\beta_1} - \frac{\alpha_2}{\beta_2}\right| \leqq \frac{|\alpha_1 - \alpha_2|}{|\beta_1|} + |\alpha_1|\cdot\left|\frac{1}{\beta_1} - \frac{1}{\beta_2}\right| \tag{1}$$
$$\leqq |\beta_1|^{-1} \, [|\alpha_1 - \alpha_2| + |\alpha_1|\cdot|\beta_2|^{-1} \, |\beta_1 - \beta_2|].$$

Next, from Lemma 1,

$$||\bar{g}^1 - \bar{g}^2|| \leqq M_B \, ||\bar{w}^1 - \bar{w}^2||,$$

(where $\bar{g}^1 = B(\bar{w}^1)$ and $\bar{g}^2 = B(\bar{w}^2)$). Thus,

$$|| \, |\bar{r}^1| - |\bar{r}^2| \, || = || \, |\bar{g}^1 - f| - |\bar{g}^2 - f| \, ||$$
$$\leqq ||\bar{g}^1 - \bar{g}^2|| \leqq M_B \, ||\bar{w}^1 - \bar{w}^2||.$$

Let $\bar{\bar{w}}^1 = F(\bar{w}^1)$ and $\bar{\bar{w}}^2 = F(\bar{w}^2)$. Then, for $i = 1, \cdots, m$,

$$\bar{\bar{w}}_i^1 = \bar{w}_i^1\cdot|\bar{r}_i^1| \Big/ \sum_{j=1}^{m} \bar{w}_j^1 \, |\bar{r}_j^1| \quad \text{and} \quad \bar{\bar{w}}_i^2 = \bar{w}_i^2\cdot|\bar{r}_i^2| \Big/ \sum_{j=1}^{m} \bar{w}_j^2 \, |\bar{r}_j^2|.$$

We have, for each i,

$$|\bar{w}_i^1\cdot|\bar{r}_i^1| - \bar{w}_i^2 \, |\bar{r}_i^2|| \leqq \bar{w}_i^1 \, ||\bar{r}_i^1| - |\bar{r}_i^2|| + |\bar{w}_i^1 - \bar{w}_i^2|\cdot|\bar{r}_i^2|$$
$$\leqq 1\cdot M_B \, ||\bar{w}^1 - \bar{w}^2|| + ||\bar{r}^2||\cdot||\bar{w}^1 - \bar{w}^2||.$$

From the equivalence of norms on finite-dimensional vector spaces, there is an $e > 0$ such that $e \, ||\cdot|| \leqq ||\cdot||_2$ (where $||\cdot||_2$ denotes the $\bar{w}^2$-weighted L_2 norm), and from the compactness of W_ϵ, there is an appropriate e which serves uniformly for every $\bar{w}^2 \in W_\epsilon$. Thus,

$$e \, ||\bar{r}^2|| \leqq ||\bar{r}^2||_2 = ||\bar{g}^2 - f||_2 \leqq ||0 - f||_2 = ||f||_2.$$

Similarly, $||f||_2$ may be uniformly bounded for every $\bar{w}^2 \in W_\epsilon$ which implies the existence of a constant M_1 such that

$$|w_i^1 \, |\bar{r}_i^1| - \bar{w}_i^2 \, |\bar{r}_i^2|| \leqq M_1 \, ||\bar{w}^1 - \bar{w}^2||.$$

This yields

$$\left|\sum_{j=1}^{m} \bar{w}_j^1 \, |\bar{r}_j^1| - \sum_{j=1}^{m} \bar{w}_j^1 \, |\bar{r}_j^1|\right| \leqq m\cdot M_1 \, ||\bar{w}^1 - \bar{w}^2||.$$

Now, applying the inequality (1) with $\alpha_1 = \bar{w}_i^1 \, |\bar{r}_i^1|$, $\alpha_2 = \bar{w}_i^2 \, |\bar{r}_i^2|$, $\beta_1 = \sum_{j=1}^{m} \bar{w}_j^1 \, |\bar{r}_j^1|$ and $\beta_2 = \sum_{j=1}^{m} \bar{w}_j^2 \, |\bar{r}_j^2|$, we have

$$|\bar{w}_i^1 - \bar{w}_i^2| \leqq \left(\sum_j \bar{w}_j^1 |\bar{r}_j^1|\right)^{-1} \Big[|\bar{w}_i^1 |\bar{r}_i^1| - \bar{w}_i^2 |\bar{r}_i^2|| + \bar{w}_i^1 |\bar{r}_i^1|$$

$$\cdot \left(\sum_j \bar{w}_j^2 |\bar{r}_j^2|\right)^{-1} \left| \sum_j \bar{w}_j^1 |\bar{r}_j^1| - \sum_j \bar{w}_j^2 |\bar{r}_j^2| \right| \Big]$$

$$\leqq M_F \, ||\bar{w}^1 - \bar{w}^2||$$

for some M_F depending only upon ϵ (the uniform bound on $(\sum_i w_i |r_i|)^{-1}$ for $w \in W$, follows as in similar cases before). □

LEMMA 3. *If, for any j, $|g^*(x_j) - f(x_j)| < \tau^*$ then $w_j^k \to 0$ as $k \to \infty$.*

Proof. The sequence $\{w_j^k\}_{k=0}^\infty$, being bounded above and below by 1 and 0 respectively, has a convergent subsequence. To prove the lemma, it suffices to show every convergent subsequence has limit zero. To this end, suppose $\{w_j^{k_i}\}_{i=0}^\infty$ is a convergent subsequence with limit $a > 0$. Select N such that $i \geqq N$ implies $w_j^{k_i} > a/2$, then

$$\sigma_{k_i}^2 = \sum_{l=1}^m w_l^{k_i} |\bar{g}^{k_i}(x_l) - f(x_l)|^2 \leqq \sum_{l=1}^m w_l^{k_i} |g^*(x_l) - f(x_l)|^2$$

$$\leqq w_j^{k_i} |g^*(x_j) - f(x_j)|^2 + \sum_{l \neq j} w^{k_i} |g^*(x_l) - f(x_l)|^2$$

$$\leqq w_j^{k_i} |g^*(x_j) - f(x_j)|^2 + (1 - w_j^{k_i})\tau^{*2}$$

$$\leqq \tau^{*2} - w_j^{k_i}(\tau^{*2} - |g^*(x_j) - f(x_j)|^2)$$

$$\leqq \tau^{*2} - (a/2)(\tau^{*2} - |g^*(x_j) - f(x_j)|^2)$$

which implies $\{\sigma_{k_i}\}_{i=0}^\infty$ cannot have limit τ^*, contrary to assumption. □

Let

$$J_E = \{j : x_j \in E\} \quad \text{and} \quad J_0 = \{j : x_j \notin E\} = X \sim J_E.$$

LEMMA 4.

$$\sum_{j=1}^m w_j^k |\bar{g}^k(x_j) - f(x_j)| \to \tau^* \quad \text{as } k \to \infty.$$

Proof. From Lemma 3 we have that $\sum_{J_0} w_j^k \to 0$ and thus $\sum_{J_E} w_j^k \to 1$. It follows that

$$\sum_{j=1}^m w_j^k |g^*(x_j) - f(x_j)|$$

$$= \sum_{J_0} w_j^k |g^*(x_j) - f(x_j)| + \sum_{J_E} w_j^k |g^*(x_j) - f(x_j)|$$

$$= \sum_{J_0} w_j^k |g^*(x_j) - f(x_j)| + \tau^* \cdot \sum_{J_E} w_j^k$$

$$\to 0 + \tau^* \quad \text{as } k \to \infty.$$

But also

$$\left|\sum_{j=1}^{m} w_j^k |g^*(x_j) - f(x_j)| - \sum_{j=1}^{m} w_j^k |\tilde{g}^k(x_j) - f(x_j)|\right|$$

$$= \left|\sum_{j=1}^{m} w_j^k(|g^*(x_j) - f(x_j)| - |\tilde{g}^k(x_j) - f(x_j)|)\right|$$

$$\leqq \sum_{j=1}^{m} w_j^k |g^*(x_j) - \tilde{g}^k(x_j)| \leqq ||g^* - \tilde{g}^k||\cdot 1$$

$$\to 0 \quad \text{as } k \to \infty,$$

thus,

$$\sum_{j=1}^{m} w_j^k |\tilde{g}^k(x_j) - f(x_j)| \to \tau^*. \quad \square$$

LEMMA 5. *For any j, let* $\lambda_j = |g^*(x_j) - f(x_j)|/\tau^*$. *Then, either*
(i) *for some* N_j, $w_j^k = 0$ *for* $k \geqq N_j$, *or*
(ii) *for any* $\lambda' < \lambda_j$ $(\lambda' > 0)$ *and* $\lambda'' > \lambda_j$, *there exist positive constants* M' *and* M'' *such that*

$$M'\lambda'^k < w_j^k < M''\lambda''^k \quad \textit{for all } k.$$

Proof. Assume (i) does not hold, then $w_j^k > 0$ for all k. Since

$$w_j^{k+1} = w_j^k |\tilde{g}^k(x_j) - f(x_j)| \Big/ \sum_{i=1}^{m} w_i^k |\tilde{g}^k(x_i) - f(x_i)|,$$

$$\frac{w_j^{k+1}}{w_j^k} = \frac{|\tilde{g}^k(x_j) - f(x_j)|}{\sum_{i=1}^m w_i^k |\tilde{g}^k(x_i) - f(x_i)|} \to \frac{|g^*(x_j) - f(x_j)|}{\tau^*} = \lambda_j$$

as $k \to \infty$. Because $\lambda_j \in (\lambda', \lambda'')$, there is an N such that $k \geqq N$ implies $\lambda' < w_j^{k+1}/w_j^k < \lambda''$, hence, $\lambda' w_j^k < w_j^{k+1} < \lambda'' w_j^k$, and

$$\lambda'^{l+k}(\lambda'^{-k} w_j^k) = \lambda'^l w_j^k < \cdots < w_j^{k+l} < \cdots < \lambda''^l w_j^k = \lambda''^{k+l}(\lambda''^{-k} w_j^k)$$

for $l \geqq 1$. In particular, for $k = N$ and letting $i = k + l$,

$$[\lambda'^{-N} w_j^N]\lambda'^i < w_j^i < [\lambda''^{-N} w_j^N]\cdot\lambda''^i.$$

So the lemma holds for $i \geqq N$ with

$$M' = [\lambda'^{-N} w_j^N] \quad \text{and} \quad M'' = [\lambda''^{-N} w_j^N],$$

and holds for all i if M' and M'' are appropriately altered. $\square$

THEOREM 1. *For any weight* $w \in W$ *such that* $w_j = 0$ *for all* $j \in J_0$, *let* $\tilde{g} = B(w)$ *and* $r = f - \tilde{g}$. *If*

$$\operatorname{sgn} r_j = \operatorname{sgn} r_j^*, \quad \textit{for all } j \in J_E,$$

then $B(F(w)) = g^*$.

Proof. By the orthogonality property, $\tilde{g}' = B(F(w))$ is the unique element of G satisfying

$$\sum_{j=1}^{m} w_j'(f(x_j) - \tilde{g}'(x_j))\cdot g(x) = 0 \quad \text{for all } g \in G,$$

where $w' = F(w)$.

For $j \in J_E$, we have

$$|r_j| = \operatorname{sgn} r_j \cdot r_j = \operatorname{sgn} r_j^* \cdot r_j = \tau^* \cdot r_j / r_j^*$$

and thus,

$$w_j' = \alpha^{-1} w_j \ |r_j| = (\tau^*/\alpha) \cdot w_j r_j / r_j^*,$$

where $\alpha = \sum_{i=1}^m w_i \ |r_i|$. Hence, for each $g \in G$,

$$\sum_{j=1}^m w_j'(f(x_j) - g^*(x_j)) \cdot g(x_j) = \sum_{j \in J_E} w_j' r_j^* \cdot g(x_j) = (\tau^*/\alpha) \cdot \sum_{j \in J_E} w_j r_j \cdot g(x_j)$$

$$= (\tau^*/\alpha) \cdot \sum_{j=1}^m w_j r_j \cdot g(x_j) = 0$$

which implies $g^* = B(w') = B(F(w))$. □

COROLLARY 1. *If, for some k, $w_j^k = 0$ for all $j \in J_0$, and* $\operatorname{sgn} r_j^k = \operatorname{sgn} r_j^*$ *for all $j \in J_E$, then*

$$\bar{g}^{k_1} = g^* \quad \textit{for all } k_1 \geqq k + 1$$

(i.e., the algorithm converges in $k + 1$ iterations).

Proof. According to Lawson [2, p. 72], if $\sigma^{k+1} = \tau^*$, then $\bar{g}^{k_1} = g^*$ for $k_1 \geqq k + 1$. But from Theorem 1, $\bar{g}^{k+1} = g^*$ and

$$\sigma^{k+1} = \left(\sum_{j=1}^m w_j^k (f(x_j) - g^*(x_j))^2 \right)^{1/2}$$

$$= \left(\sum_{j \in J_E} w_j^k \cdot \tau^{*2} \right)^{1/2} = \tau^*. \quad \square$$

COROLLARY 2. *If, for some N, $w_j^N = 0$ for all $j \in J_0$, the algorithm converges in a finite number of steps.*

Proof. It is clear that, for $k \geqq N$, $w_j^k = 0$ for all $j \in J_0$. Determine $N_1 \geqq N$ such that $k \geqq N_1$ implies $||\bar{g}^k - g^*|| < \tau^*$. Then,

$$||r^k - r^*|| = ||(f - \bar{g}^k) - (f - g^*)|| < \tau^*$$

and hence, since $|r_j^*| = \tau^*$ for $j \in J_E$,

$$\operatorname{sgn} r_j^k = \operatorname{sgn} r_j^* \quad \text{for } j \in J_E.$$

The first corollary then guarantees that $\bar{g}^k = g^*$ for $k \geqq N_1 + 1$. □

Let J_1 be defined by

$$J_1 = \{j : j \in J_0 \text{ and } w_j^k > 0 \text{ for every } j\}$$

and let

$$\rho_0 = \max_{j \in J_1} |f(x_j) - g^*(x_j)| / \tau^*.$$

THEOREM 2. *Given any $\lambda > \rho_0$, there exists a constant M such that $||\bar{g}^k - g^*|| < M\lambda^k$ and $\tau^k - \tau^* < M\lambda^k$ for all k.*

Proof. If, for any k, $\sum_{j \in J_E} w_j^k = 0$, then for all greater k the same is true which violates Lemma 4. Letting $\beta_k = \sum_{j \in J_E} w_j^k$, we may define a new weight $\bar{w}^k$ for each k by

$$\bar{w}_j^k = 0, \qquad j \in J_0,$$
$$= \beta_k^{-1} \cdot w_j^k, \qquad j \in J_E.$$

From Lemma 5, we may assert the existence of a constant M such that for $j \in J_1$, $|\bar{w}_j^k - w_j^k| = w_j^k < M\lambda^k$ since $\lambda > \lambda_j$ for all $j \in J_1$.

For $j \in J_E$,

$$|\bar{w}_j^k - w_j^k| = (\beta_k^{-1} - 1) \cdot w_j^k \leqq \beta_k^{-1} - 1 = \beta_k^{-1} \sum_{i \in J_0} w_i^k \leqq \beta_k^{-1} \cdot m \cdot M\lambda^k.$$

But since $\beta_k \to 1$ as $k \to \infty$, $\{\beta_k^{-1}\}$ is uniformly bounded from above and we may assert the existence of an M_1 such that

$$||\bar{w}^k - w^k|| < M_1\lambda^k \quad \text{for all } k.$$

It is now claimed that, for some $\epsilon > 0$, both of the sequences $\{w^k\}_{k=0}^{\infty}$ and $\{\bar{w}^k\}_{k=0}^{\infty}$ are contained in the set W_ϵ. Considering $\{w^k\}_{k=0}^{\infty}$ first, if this sequence is not contained entirely in any such W_ϵ there exists a convergent subsequence with limit w^* such that $w_j^* > 0$ for at most $n - 1$ values of j, thus $w^* \notin W$, contrary to the fact that every limit point of $\{w^k\}_{k=0}^{\infty}$ is an element of W (see Lawson [2, pp. 75–76]). Thus, $\{w^k\}_{k=0}^{\infty}$ is contained in W_ϵ for some positive ϵ; that the same is true for $\{\bar{w}^k\}_{k=0}^{\infty}$ follows from the same argument and the fact that $||w^k - \bar{w}^k|| \to 0$ as $k \to \infty$.

We may now apply Lemma 2 to guarantee the existence of an M_2 such that

$$||F(\bar{w}^k) - w^{k+1}|| = ||F(\bar{w}^k) - F(w^k)|| \leqq M_2\, ||\bar{w}^k - w^k|| \leqq M_2 \cdot M_1\lambda^k.$$

We now define $\bar{g}^k$ to be $B(\bar{w}^k)$ and hence from Lemma 1, there is an M_3 such that

$$||g^k - \bar{g}^k|| = ||B(\bar{w}^k) - B(w^k)|| \leqq M_3\, ||\bar{w}^k - w^k|| \leqq M_3\lambda^k.$$

Select N so large that $||g^k - \bar{g}^k|| < \tau^*/2$ and $||g^* - g^k|| < \tau^*/2$ for $k \geqq N$. Hence, $||g^* - \bar{g}^k|| < \tau^*$ and as in the proof to the second corollary of Theorem 1, $\text{sgn } r_j^* = \text{sgn } \bar{r}_j^k$ for all $j \in J_E$.

Applying Theorem 1 to $\bar{w}^k$, we see that

$$B(F(\bar{w}^k)) = g^* \quad \text{for } k \geqq N.$$

Applying Lemma 1 again, we have

$$||g^* - g^{k+1}|| = ||B(F(\bar{w}^k)) - B(w^{k+1})|| \leqq M_4\lambda^{k+1} \quad \text{for } k \geqq N$$

and suitable M_4. Therefore, the inequality holds for all k with larger M_4 if necessary. The proof is completed by noticing that

$$\tau_k - \tau^* = ||f - g^k|| - ||f - g^*|| \leqq ||g^k - g^*||. \quad \square$$

Several closing comments would be instructive. First, since

$$\rho = \max_{j \in J_0} |f(x_j) - g^*(x_j)|/\tau^*,$$

it is clear that $\rho_0 \leqq \rho$, hence, Theorem 2 also holds for ρ. If $\rho_0 < \rho$, then, for some $l \in J_0$,

$$|f(x_l) - g^*(x_l)| \geqq |f(x_j) - g^*(x_j)|$$

for all $j \in J_0$, and, for some N, $w_i^k = 0$ for $k \geqq N$. Computationally, this is unlikely with the standard algorithm. Thus, the convergence factor may be assumed to be ρ.

Several techniques to accelerate the convergence of Lawson's algorithm have been tried (see Rice and Usow [4] and Cline [1, pp. 103–121]). The most successful techniques involve monitoring the quantities w_j^k and setting to zero those for which very probably $j \in J_0$. Usually, these are j such that $|f(x_j) - g^*(x_j)|/\tau^*$ is very small, and hence less than ρ_0. If we assume $\tau^k - \tau^* \approx M\rho_0^k$, then altering such w_j^k will have no effect on the asymptotic behavior of $\{\tau^k - \tau^*\}$, but may on the initial behavior. This has been observed in numerical experiments. To decrease the asymptotic rate, hence ρ_0, it would be necessary to set to zero w_j^k where $|f(x_j) - g^*(x_j)|/\tau^*$ is less than 1 but very close to 1. This is extremely difficult since $|f - g^*|$ is only known approximately as $|f - \tilde{g}^k|$.

Computing Facility
National Center for Atmospheric Research*
P. O. Box 1470
Boulder, Colorado 80302

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

1. A. K. Cline, *Uniform Approximation as a Limit of L_2 Approximations*, Ph.D. Thesis, University of Michigan, Ann Arbor, Mich., 1970.
2. C. L. Lawson, *Contributions to the Theory of Linear Least Maximum Approximation*, Ph.D. Thesis, University of California, Los Angeles, Calif., 1961.
3. J. R. Rice, *The Approximation of Functions*. Vol. II: *Nonlinear and Multivariate Theory*, Addison-Wesley, Reading, Mass., 1969, pp. 298–304. MR **39** #5989.
4. J. R. Rice & K. H. Usow, "The Lawson algorithm and extensions," *Math. Comp.*, v. 22, 1968, pp. 118–127. MR **38** #463.