SIAM J. NUMER. ANAL.
Vol. 40, No. 2, pp. 792–804

A CONVERGENCE ANALYSIS OF DYKSTRA'S ALGORITHM FOR POLYHEDRAL SETS*

CHRIS PERKINS†

Abstract. Let H be a nonempty closed convex set in a Hilbert space X determined by the intersection of a finite number of closed half spaces. It is well known that given an $x_0 \in X$, Dykstra's algorithm applied to this collection of closed half spaces generates a sequence of iterates that converge to $P_H(x_0)$, the orthogonal projection of x_0 onto H. The iterates, however, do not necessarily lie in H. We propose a combined Dykstra–conjugate-gradient method such that, given an $\varepsilon > 0$, the algorithm computes an $x \in H$ with $\|x - P_H(x_0)\| < \varepsilon$. Moreover, for each iterate x_m of Dykstra's algorithm we calculate a bound for $\|x_m - P_H(x_0)\|$ that approaches zero as m tends to infinity. Applications are made to computing bounds for $\|x_m - P_H(x_0)\|$ where H is a polyhedral cone. Numerical results are presented from a sample isotone regression problem.

Key words. Dykstra's algorithm, best approximation from polyhedra, Hildreth's algorithm, linear inequalities

AMS subject classifications. Primary, 41A65; Secondary, 65F10, 15A39

PII. S0036142900367557

1. Introduction.

Let X be a Hilbert space and $K_0, K_1, \ldots, K_{n-1}$ be n closed convex sets in X whose intersection, K, is nonempty. Given an $x_0 \in X$, the Boyle–Dykstra algorithm [3] generates a sequence of iterates $\{x_m\}$ whose limit is the orthogonal projection of x_0 onto K. Defining P_{K_i} to be the orthogonal projection onto K_i, letting $[m]$ denote m mod n, and setting

$$e_{-n} = e_{-(n-1)} = \cdots = e_{-1} = \vec{0},$$

Dykstra's algorithm can be written as

$$\begin{aligned} x_{m+1} &= P_{K_{[m]}}(x_m + e_{m-n}), \\ e_m &= (x_m + e_{m-n}) - x_{m+1} \end{aligned}$$

for $m = 0, 1, \ldots$. If the n closed convex sets are half spaces of the form $H_i = \{x \mid \langle x, z_i \rangle \leq f_i\}$, where $z_i \in X$ and $f_i \in \Re$, then the formula for the $(m+1)$st iterate of Dykstra's algorithm can be written as $x_{m+1} = x_m + \xi_m z_{[m]}$, where $\xi_m \in \Re$.

For this finite collection of intersecting closed half spaces, Iusem and De Pierro [11] showed that the Dykstra algorithm converged linearly. Subsequently Deutsch and Hundal [6] have sharpened the rate of convergence bound established in [11] and have shown that

$$\|x_m - P_H(x_0)\| \leq \rho c^m, \tag{1.1}$$

where $H = \cap_{i=0}^{n-1} H_i$, ρ is a constant, $0 \leq c < 1$, and $m = 0, 1, \ldots$. In general, the use of the Deutsch–Hundal result to explicitly bound $\|x_m - P_H(x_0)\|$ is problematic in that although c can be calculated a priori, ρ cannot be. This means that in practice when computing x_m, inequality (1.1) cannot be used to calculate a bound for $\|x_m - P_H(x_0)\|$

*Received by the editors January 10, 2000; accepted for publication (in revised form) October 22, 2001; published electronically July 24, 2002.
http://www.siam.org/journals/sinum/40-2/36755.html

†13656 Spring Mill Boulevard, Carmel, IN 46032 (christopher.perkins@roche.com).

for $m = 0, 1, \ldots$. Xu comments in [15] that the applicability of Dykstra's algorithm for polyhedron approximation is restrictive unless it is possible to find an active bound for ρ. This lack of a computable error bound for the Dykstra iterates is the motivation for this manuscript. Here we develop a method to explicitly bound $\|x_m - P_H(x_0)\|$ by a ρ_m, where $\lim_{m\to\infty} \rho_m = 0$. This convergence analysis allows the algorithm to be terminated when x_m is within a prescribed distance of $P_H(x_0)$.

In section 2, we first review the equivalence between Dykstra's algorithm for polyhedral sets in the general Hilbert space setting and the finite dimensional Hildreth algorithm [10]. This equivalence is critical to our computation of the error bounds for the iterates. The finite dimensional version allows us to use a linear algebraic analysis to determine a ρ_m such that $\|x_m - P_H(x_0)\| \le \rho_m$ with $\lim_{m\to\infty} \rho_m = 0$. Given that it is possible to determine when an x_m is arbitrarily close to $P_H(x_0)$, we are then able to calculate an $x \in H$ such that $\|x - P_H(x_0)\| < \varepsilon$.

Section 3 shows how the previously established results can be modified if $\{z_i\}$ is linearly independent. Other implementation details are discussed and the N-convex regression problems are reviewed. Numerical results are presented from a sample isotone regression problem.

There are other methods available for computing $P_H(x_0)$. These include solving a related least distance quadratic programming problem [7]. More recently, Xu has published an algorithm with an error analysis that also computes the nearest point mapping for polyhedrons in a finite number of iterations [15]. Unfortunately, the use of Xu's algorithm requires a priori knowledge of a point in H. Such a point can first be calculated using the algorithm proposed by T. S. Motzkin [14] and discussed in Agmon [1].

2. Definitions, lemmas, and theory. In 1988 Han [9] rediscovered Hildreth's algorithm [10] and commented that the arguments used to show convergence in Euclidean space can be used in the general Hilbert space setting where the intersecting closed convex sets are half spaces. Deutsch and Hundal [6] later reported that in Euclidean space, where the intersection of closed convex sets in polyhedral, Dykstra's algorithm reduces to Hildreth's algorithm.

In fact, the Deutsch–Hundal statement can be strengthened, and Han's result is immediately applicable to the general Hilbert space setting. This is due to the fact that in Hilbert space, Dykstra's algorithm applied to a finite set of intersecting closed half spaces is equivalent to Hildreth's algorithm applied to a related set of linear inequalities. We next review this equivalence.

Let X be a Hilbert space and $H_0, H_1, \ldots, H_{n-1}$ be closed half spaces of the form $H_i = \{x \mid \langle x, z_i\rangle \le f_i\}$ for $0 \le i \le n-1$, where $z_i \in X$, $\|z_i\| = 1$, $f_i \in \Re$, and $H = \cap_{i=0}^{n-1} H_i$ is nonempty. The boundary of each closed half space H_i is the hyperplane B_i, where $B_i = \{x \mid \langle x, z_i\rangle = f_i\}$.

LEMMA 2.1. *Given $x \in X$,*

$$P_{H_i}(x) = \begin{cases} x & \text{if } x \in H_i, \\ x - (\langle x, z_i\rangle - f_i) z_i & \text{if } x \notin H_i, \end{cases}$$

$$P_{B_i}(x) = x - (\langle x, z_i\rangle - f_i) z_i,$$

$$\mathrm{d}(x, B_i) = |\langle x, z_i\rangle - f_i|$$

for $i = 0, 1, \ldots, n-1$. □

LEMMA 2.2. *Given an $x_0 \in X$, the iterates of Dykstra's algorithm lie in* $\mathrm{span}\{x_0, z_0, \ldots, z_{n-1}\}$.

Proof. This follows from Dykstra's algorithm, the fact that the closed convex sets are half spaces, Lemma 2.1, and a simple inductive argument. □

By Lemma 2.2, the Dykstra iterates $\{x_m\}$ are contained in a finite dimensional subspace of X. Thus we can use Gram–Schmidt to generate an orthonormal basis $\{w_j\}_{j=1}^k$ for span$\{x_0, z_0, \ldots, z_{n-1}\}$, and we can write

$$x_0 = \sum_{j=1}^{k} \gamma_j w_j \tag{2.1}$$

as well as

$$z_i = \sum_{j=1}^{k} \alpha_{ij} w_j \tag{2.2}$$

for $0 \le i \le n-1$.

Since any finite dimensional Hilbert space is isometric to $l_2(k)$, we reduce the convergence analysis of Dykstra's algorithm in X to an equivalent analysis in $\Re^k$.

Defining

$$y_0 = (\gamma_1, \ldots, \gamma_k)^{\mathrm{T}} \tag{2.3}$$

and

$$a_i = (\alpha_{i1}, \ldots, \alpha_{ik})^{\mathrm{T}} \tag{2.4}$$

for $0 \le i \le n-1$, it is possible to define the n closed half spaces in $\Re^k$ by

$$h_i = \{y \mid \langle y, a_i \rangle \le f_i\} \tag{2.5}$$

and the corresponding boundaries by

$$b_i = \{y \mid \langle y, a_i \rangle = f_i\}. \tag{2.6}$$

We remark that Dykstra's algorithm applied to $h_0, h_1, \ldots, h_{n-1}$ is the same as Hildreth's algorithm applied to $Ay \le f$, where $A^{\mathrm{T}} = [a_0, \ldots, a_{n-1}]$ and $f = (f_0, \ldots, f_{n-1})^{\mathrm{T}}$.

Throughout what follows, x_0 will be defined as in (2.1), $\{x_m\}$ will denote the sequence of Dykstra iterates generated by the closed half spaces $H_0, \ldots, H_{n-1}$ in X. Similarly, y_0 will be defined as in (2.3) and $\{y_m\}$ will denote the sequence of Dykstra iterates generated by the closed half spaces $h_0, \ldots, h_{n-1}$ in $\Re^k$ as specified in (2.5).

Let $S \subset \{0, 1, \ldots, n-1\}$. We define $\tilde{S} = \{i \mid 0 \le i \le n-1 \text{ and } i \notin S\}$ and $|S|$ to be the cardinality of set S. We introduce the notation H_S (resp., h_S) to denote $\cap_{i \in S} H_i$ (resp., $\cap_{i\in S} h_i$) and B_S (resp., b_S) to denote $\cap_{i\in S} B_i$ (resp., $\cap_{i\in S} b_i$). For notational simplicity, if $S = \{0, 1, \ldots, n-1\}$, the subscript S is omitted. A_S^{T} and f_S are defined to be

$$A_S^{\mathrm{T}} = [a_{s_1}, \ldots, a_{s_{|S|}}\}, \tag{2.7}$$

$$f_S = (f_{s_1}, \ldots, f_{s_{|S|}})^{\mathrm{T}}, \tag{2.8}$$

where $s_j \in S$, $1 \le j \le |S|$ with $s_j < s_{j+1}$. Let $C \subset \{0, 1, \ldots, n-1\}$ contain the indices of the critical boundaries. That is, $C = \{i \mid P_h(y_0) \in b_i\}$. If $C = \emptyset$, then $P_h(y_0) \in \mathrm{int}(h)$ and $P_h(y_0) = y_0$. Henceforth we will assume that $C \ne \emptyset$.

Let $S \subset \{0, 1, \ldots, n-1\}$, $S \neq \emptyset$, and let A_S^{T} and f_S be as defined in (2.7) and (2.8). We denote the residual $r_{m,S} = f_S - A_S y_m$. If $f_S \in \mathrm{Range}(A_S)$ (equivalently $b_S \neq \emptyset$), then by (2.6),

$$\begin{aligned} r_{m,S} &= A_S P_{b_S}(y_m) - A_S y_m, \\ &= A_S(P_{b_S}(y_m) - y_m). \end{aligned}$$

A_S^+ is the Moore–Penrose generalized inverse of A_S, and $A_S^+ A_S$ is the projection onto the orthogonal complement of the null space of A_S (denoted $\mathrm{N}(A_S)^\perp$) [8]. If $b_S \neq \emptyset$, then $P_{b_S}(y_m) - y_m$ is an element of $\mathrm{N}(A_S)^\perp$,

$$A_S^+ A_S(P_{b_S}(y_m) - y_m) = P_{b_S}(y_m) - y_m,$$

and $A_S^\perp r_{m,S} = P_{b_S}(y_m) - y_m$. Therefore,

$$\|y_m - P_{b_S}(y_m)\| = \|A_S^+ r_{m,S}\| \leq \|A_S^+\| \, \|r_{m,S}\| \tag{2.9}$$

with $\|A_S^+\|$ being the inverse of the smallest nonzero singular value of A_S [8].

For each iterate y_m of Dykstra's algorithm we determine the existence of a set C_m with the properties

$$f_{C_m} \in \mathrm{Range}(A_{C_m}), \tag{2.10a}$$

$$\text{for all } j \in \tilde{C}_m, \quad 2\|A_{C_m}^+\| \, \|r_{m,C_m}\| < f_j - \langle y_m, a_j \rangle, \tag{2.10b}$$

$$\begin{aligned} &\text{for all } j \in \tilde{C}_m, \quad e_{\pi_{m,j}} = \vec{0}, \\ &\text{where } \pi_{m,j} \in \{i \mid m-n \leq i \leq m-1 \text{ and } [i] = j\}. \end{aligned} \tag{2.10c}$$

If no such C_m exists, we define the bound, ρ_m, on $\|y_m - P_h(y_0)\|$ to be ∞. In Theorem 2.16 we prove that for sufficiently large m, the set C satisfies properties (2.10a)–(2.10c).

For notational simplicity throughout what follows, we will use the notation r_m in place of r_{m,C_m} or $r_{m,C}$ where appropriate. We next assume that $C_m \neq \emptyset$ and develop the theory to bound $\|y_m - P_h(y_0)\|$ by ρ_m, where

$$\rho_m = 2\|A_{C_m}^+\| \, \|r_m\|. \tag{2.11}$$

Thereafter, in Theorem 2.17 we will show sufficient conditions on the selection of $\{C_m\}$ to guarantee the bound $\rho_m = 2\|A_{C_m}^+\| \, \|r_m\|$ tends to zero.

By the construction of C_m and inequality (2.9),

$$\begin{aligned} 2\|y_m - P_{b_{C_m}}(y_m)\| &\leq 2\|A_{C_m}^+\| \, \|r_m\| && (2.12) \\ &< f_j - \langle y_m, a_j \rangle && (2.13) \\ &= \mathrm{d}(y_m, b_j) \end{aligned}$$

for all $j \in \tilde{C}_m$.

The following lemma shows that if $C_m \neq \{0, 1, \ldots, n-1\}$, then the closed ball centered at $P_{b_{C_m}}(y_m)$ with radius $\mathrm{d}(y_m, P_{b_{C_m}}(y_m))$ is contained in the $\mathrm{int}(h_{\tilde{C}_m})$. To facilitate subsequent discussion, we define

$$\mathcal{B}_m = \mathcal{B}[P_{b_{C_m}}(y_m), \mathrm{d}(y_m, P_{B_{C_m}}(y_m))]. \tag{2.14}$$

LEMMA 2.3. *If $C_m \neq \{0, 1, \dots, n-1\}$, then the closed ball $\mathcal{B}_m \subset \mathrm{int}(h_{\tilde{C}_m})$.*

Proof. We will show that, given any element y contained in the closed ball $\mathcal{B}_m$, $\langle y, a_j \rangle < f_j$ for every $j \in \tilde{C}_m$. This will imply that $y \in \mathrm{int}(h_{\tilde{C}_m})$. Let $y \in \mathcal{B}_m$ and define $\varepsilon_1 = \mathrm{d}(y_m, P_{b_{C_m}}(y_m))$. Then $P_{b_{C_m}}(y_m) = y_m + \varepsilon_1 v_1$ for some unit vector v_1. Since $y \in \mathcal{B}_m$, $y = P_{b_{C_m}}(y_m) + \varepsilon_2 v_2$ for $0 \leq \varepsilon_2 \leq \mathrm{d}(y_m, P_{b_{C_m}}(y_m)) = \varepsilon_1$ and some unit vector v_2. Let $j \in \tilde{C}_m$. Then

$$\begin{aligned}
\langle y, a_j \rangle &= \langle P_{b_{C_m}}(y_m) + \varepsilon_2 v_2, a_j \rangle \\
&= \langle y_m + \varepsilon_1 v_1 + \varepsilon_2 v_2, a_j \rangle \\
&= \langle y_m, a_j \rangle + \varepsilon_1 \langle v_1, a_j \rangle + \varepsilon_2 \langle v_2, a_j \rangle \\
&\leq \langle y_m, a_j \rangle + \varepsilon_1 \|v_1\| \, \|a_j\| + \varepsilon_2 \|v_2\| \, \|a_j\| \\
&= \langle y_m, a_j \rangle + \varepsilon_1 + \varepsilon_2 \\
&\leq \langle y_m, a_j \rangle + 2\varepsilon_1 \\
&= \langle y_m, a_j \rangle + 2\mathrm{d}(y_m, P_{b_{C_m}}(y_m)) \\
&= \langle y_m, a_j \rangle + 2\|y_m - P_{b_{C_m}}(y_m)\| \\
&< \langle y_m, a_j \rangle + f_j - \langle y_m, a_j \rangle \quad \text{by (2.13)} \\
&= f_j.
\end{aligned}$$

Thus $y \in \mathrm{int}(h_j)$. Since j was arbitrary in $\tilde{C}_m$, $y \in \mathrm{int}(h_{\tilde{C}_m})$, and $\mathcal{B}_m \subset \mathrm{int}(h_{\tilde{C}_m})$. ∎

In order to show the bound stated in (2.11), we will prove that all subsequent Dykstra iterates are elements of the closed ball $\mathcal{B}_m$. We will then have $\{y_{m+k}\}_{k=0}^{\infty} \subset \mathcal{B}_m$. $\mathcal{B}_m$, by construction, is closed and must contain its limit points. Since $\{y_{m+k}\}_{k=0}^{\infty}$ converges to $P_h(y_0)$, $P_h(y_0)$ would then be contained in $\mathcal{B}_m$. This will allow us to compute a bound on $\|y_m - P_h(y_0)\|$.

The following result is proved in [6]. As in section 1, $[m]$ denotes m mod n.

RESULT 2.4. *If y_m is the mth iterate of Dykstra's algorithm, then $y_{m+1} = (1-\lambda)y_m + \lambda P_{b_{[m]}}(y_m)$ for some λ, $0 \leq \lambda \leq 1$.*

The following proposition can be proved in part by using the characterization of best approximation from subspaces in Hilbert space [5].

PROPOSITION 2.5. *Let $y \in \Re^k$, $S \subset \{0, 1, \dots, n-1\}$, $S \neq \emptyset$, $b_S \neq \emptyset$, $m \geq 0$ with $[m] \in S$; then for $0 \leq \lambda \leq 1$*

$$\|(1-\lambda)y + \lambda P_{b_{[m]}}(y) - P_{b_S}(y)\| \leq \|y - P_{b_S}(y)\|.$$

Let $S \neq \emptyset$ and define the subspace

$$\hat{b}_S = \bigcap_{i \in S} \{y \mid \langle y, a_i \rangle = 0\}.$$

If $b_S \neq \emptyset$, then $P_{b_S}(y) = P_{\hat{b}_S}(y) + P_{b_S}(\vec{0})$. Since $P_{\hat{b}_S}$ is continuous, Proposition 2.6 follows.

PROPOSITION 2.6. *Let $S \neq \emptyset$ and $b_S \neq \emptyset$; then P_{b_S} is continuous.*

PROPOSITION 2.7. *Let $y \in \Re^k$, $S \neq \emptyset$, and $b_S \neq \emptyset$; then*

$$P_{b_S}\left(y + \sum_{i \in S} \xi_i a_i\right) = P_{b_S}(y).$$

Proposition 2.7 can be shown using the fact that $P_{b_S}(y) = P_{\hat{b}_S}(y) + P_{b_S}(\vec{0})$, $P_{\hat{b}_S}$ is a linear operator, and $\hat{b}_S$ is the orthogonal complement of $\text{span}_{i \in S}\{a_i\}$.

Lemma 2.8 next shows that y_m and all subsequent Dykstra iterates are elements of the closed ball $\mathcal{B}_m$. As previously defined, $\mathcal{B}_m = \mathcal{B}[P_{b_{C_m}}(y_m), \mathrm{d}(y_m, P_{b_{C_m}}(y_m))]$. The lemma is actually proven by showing a stronger result as stated in the induction hypothesis.

LEMMA 2.8. *The iterates of Dykstra's algorithm y_{m+k}, $k = 0, 1, \ldots$, are contained in the closed ball $\mathcal{B}_m$.*

Proof. The proof is by induction.

It is evident that $y_m \in \mathcal{B}_m$. Moreover, by the definition of C_m (2.10c), for all $y \in \tilde{C}_m$, $e_{\pi_{m,j}} = \vec{0}$, where $\pi_{m,j} \in \{i \mid m - n \leq i \leq m - 1 \text{ and } [i] = j\}$.

We will assume that for each l, $1 \leq l \leq k$, $y_{m+l} \in \mathcal{B}_m$ and for all $j \in \tilde{C}_m$, $e_{\pi_{m+l,j}} = \vec{0}$, where $\pi_{m+l,j} \in \{i \mid m + l - n \leq i \leq m + l - 1 \text{ and } [i] = j\}$.

Next we will show that $y_{m+k+1} \in \mathcal{B}_m$.

If $[m+k] \in \tilde{C}_m$, then, by Lemma 2.3 and the induction hypothesis, $y_{m+k} \in h_{[m+k]}$. Therefore

$$y_{m+k+1} = P_{h_{[m+k]}}(y_{m+k} + \vec{0}) = P_{h_{[m+k]}}(y_{m+k}) = y_{m+k}.$$

Since $y_{m+k} \in \mathcal{B}_m$, $y_{m+k+1} \in \mathcal{B}_m$.

If $[m + k] \in C_m$, then by applying Result 2.4 to the $m + k + 1$ iterate,

$$y_{m+k+1} = (1 - \lambda)y_{m+k} + \lambda P_{b_{[m+k]}}(y_{m+k})$$

for some λ, $0 \leq \lambda \leq 1$. By Proposition 2.5,

$$\|(1 - \lambda)y_{m+k} + \lambda P_{b_{[m+k]}}(y_{m+k}) - P_{b_{C_m}}(y_{m+k})\| \leq \|y_{m+k} - P_{b_{C_m}}(y_{m+k})\|$$

for $0 \leq \lambda \leq 1$. Thus,

$$\|y_{m+k+1} - P_{b_{C_m}}(y_{m+k})\| \leq \|y_{m+k} - P_{b_{C_m}}(y_{m+k})\|.$$

Using the induction hypothesis, we have $y_{m+l} = y_{m+l-1}$ whenever $[m + l - 1] \in \tilde{C}_m$. Therefore, it can be shown that $y_{m+k} = y_m + \sum_{j \in C_m} \xi_j a_j$ and, by Proposition 2.7, $P_{b_{C_m}}(y_{m+k}) = P_{b_{C_m}}(y_m)$.

Therefore

$$\|y_{m+k+1} - P_{b_{C_m}}(y_m)\| \leq \|y_{m+k} - P_{b_{C_m}}(y_m)\|.$$

Since $y_{m+k} \in \mathcal{B}_m$, it follows that $y_{m+k+1} \in \mathcal{B}_m$.

Moreover, regardless of whether $[m + k] \in C_m$ or $[m + k] \in \tilde{C}_m$ for all $j \in \tilde{C}_m$, $e_{\pi_{m+k+1,j}} = \vec{0}$, where $\pi_{m+k+1,j} \in \{i \mid m + k + 1 - n \leq i \leq m + k \text{ and } [i] = j\}$.

Therefore, for $k = 0, 1, \ldots, y_{m+k} \in \mathcal{B}_m$. $\square$

We next bound $\|y_m - P_h(y_0)\|$.

LEMMA 2.9. $\|y_m - P_h(y_0)\| \leq 2\|y_m - P_{b_{C_m}}(y_m)\|$.

Proof. By Lemma 2.8, $\{y_{m+k}\}_{k=0}^{\infty} \subset \mathcal{B}_m$. Since $\mathcal{B}_m$ is closed and $\lim_{k\to\infty} y_{m+k} = P_h(y_0)$, we have that $P_h(y_0) \in \mathcal{B}_m$. By the triangle inequality,

$$\begin{aligned}\|y_m - P_h(y_0)\| &\le \|y_m - P_{b_{C_m}}(y_m)\| + \|P_{b_{C_m}}(y_m) - P_h(y_0)\| \\ &\le 2\|y_m - P_{b_{C_m}}(y_m)\|. \qquad \square\end{aligned}$$

The first major result of this section follows.

THEOREM 2.10. $\|y_m - P_h(y_0)\| \le 2\|A_{C_m}^{+}\|\ \|r_m\|$.

Proof. By inequality (2.12), $2\|y_m - P_{b_{C_m}}(y_m)\| \le 2\|A_{C_m}^{+}\|\ \|r_m\|$. By Lemma 2.9, $\|y_m - P_h(y_0)\| \le 2\|y_m - P_{b_{C_m}}(y_m)\|$ and thus $\|y_m - P_h(y_0)\| \le 2\|A_{C_m}^{+}\|\ \|r_m\|$. $\square$

We now show the prerequisites for proving Theorems 2.16 and 2.17. Lemma 2.11 shows that $C = \{i \mid P_h(y_0) \in b_i\}$ satisfies property (2.10a).

LEMMA 2.11. $f_C \in \mathrm{Range}(A_C)$.

Proof. By definition of C, $P_h(y_0) \in b_C$. Therefore $A_C P_h(y_0) = f_C$ and $f_C \in \mathrm{Range}(A_C)$. $\square$

If $C = \{0, 1, \ldots, n-1\}$, then since y_m converges to $P_h(y_0)$, there exists a smallest integer M such that, for all $m > M - n$, $\|y_m - P_h(y_0)\| < \frac{1}{3\kappa}$, where $\kappa = \|A^{+}\|\ \|A\|$. Otherwise $C \ne \{0, 1, \ldots, n-1\}$, and again by the convergence of y_m to $P_h(y_0)$, there exists a smallest integer M such that, for all $m > M - n$, $\|y_m - P_h(y_0)\| < \frac{\varepsilon_C}{3\kappa_C}$, where $\kappa_C = \|A_C^{+}\|\ \|A_C\|$ and $\varepsilon_C > 0$ such that $\mathcal{B}[P_h(y_0), \varepsilon_C] \subset \mathrm{int}(h_{\tilde{C}})$.

LEMMA 2.12. *For all $m > M$ and $j \in \tilde{C}$, $e_{\pi_{m,j}} = \vec{0}$, where $\pi_{m,j} \in \{i \mid m - n \le i \le m - 1$ and $[i] = j\}$.*

Proof. If $C = \{0, 1, \ldots, n-1\}$, then the lemma is vacuously true. Next assume that $C \ne \{0, 1, \ldots, n-1\}$ and let $m > M$. Suppose there exists an $j \in \tilde{C}$ such that $e_{\pi_{m,j}} \ne \vec{0}$, where $\pi_{m,j} \in \{i \mid m - n \le i \le m - 1$ and $[i] = j\}$. Since $e_{\pi_{m,j}} \ne \vec{0}$, this implies that $y_{\pi_{m,j}+1} = P_{b_j}(y_{\pi_{m,j}})$ and $y_{\pi_{m,j}+1} \in b_j$. This is impossible since $\pi_{m,j} + 1 > M - n$ and $y_{\pi_{m,j}+1} \in \mathcal{B}[P_h(y_0), \varepsilon_C] \subset \mathrm{int}(h_j)$. Clearly $\langle y_{\pi_{m,j}+1}, a_j\rangle < f_j$ and $\langle y_{\pi_{m,j}+1}, a_j\rangle = f_j$ cannot both be true. Therefore $e_{\pi_{m,j}} = \vec{0}$.

Thus for all $m > M$ and $j \in \tilde{C}$, $e_{\pi_{m,j}} = \vec{0}$, where $\pi_{m,j} \in \{i \mid m - n \le i \le m - 1$ and $[i] = j\}$. $\square$

Lemma 2.12 implies the following result.

RESULT 2.13. *For all $m > M$, $y_{m+1} = y_m$ whenever $[m] \in \tilde{C}$.*

We now prove that for sufficiently large m, $2\|A_C^{+}\|\ \|r_m\| < f_j - \langle y_m, a_j\rangle$ for all $j \in \tilde{C}$.

LEMMA 2.14. *If $\{y_m\}$ is the sequence of Dykstra iterates, then for all $m > M$, $P_{b_C}(y_m) = P_h(y_0)$.*

Proof. As a consequence of Result 2.13 and the definition of Dykstra's algorithm, for all $m > M$,

$$y_m = y_{M+1} + \sum_{i\in C} \xi_i a_i.$$

Using Proposition 2.7,

$$\begin{aligned}P_{b_C}(y_m) &= P_{b_C}\left(y_{M+1} + \sum_{i\in C} \xi_i a_i\right) \\ &= P_{b_C}(y_{M+1}).\end{aligned}$$

By Proposition 2.6, P_{b_C} is continuous, and thus for all $m > M$

$$\begin{aligned} P_{b_C}(y_m) &= P_{b_C}(y_{M+1}) \\ &= \lim_{m\to\infty} P_{b_C}(y_m) \\ &= P_{b_C}\left(\lim_{m\to\infty} y_m\right) \\ &= P_{b_C}(P_h(y_0)) \\ &= P_h(y_0). \end{aligned}$$

The last equality holds since $P_h(y_0) \in b_C$. ∎

As specified above, $\kappa_C = \|A_C^+\| \, \|A_C\|$, $\varepsilon_C > 0$ such that $\mathcal{B}[P_h(y_0), \varepsilon_C] \subset \operatorname{int}(h_{\tilde{C}})$, and M is selected such that for all $m > M - n$ we have $\|y_m - P_h(y_0)\| < \frac{\varepsilon_C}{3\kappa_C}$. Moreover, it can be shown that $\kappa_C \geq 1$ [12].

LEMMA 2.15. *If $m > M$, then $2\|A_C^+\| \, \|r_m\| < f_j - \langle y_m, a_j\rangle$ for all $j \in \tilde{C}$.*

Proof. If $C = \{0, 1, \ldots, n-1\}$, then the lemma is vacuously true. Next assume that $C \neq \{0, 1, \ldots, n-1\}$ and let $m > M$.

$$\begin{aligned} 2\|A_C^+\| \, \|r_m\| &= 2\|A_C^+\| \, \|f_C - A_C y_m\| \\ &= 2\|A_C^+\| \, \|A_C(P_{b_C}(y_m) - y_m)\| \\ &\leq 2\|A_C^+\| \, \|A_C\| \, \|y_m - P_{b_C}(y_m)\| \\ &\leq 2\kappa_C \|y_m - P_{b_C}(y_m)\| \\ &= 2\kappa_C \|y_m - P_h(y_0)\| \\ &< 2\kappa_C \frac{\varepsilon_C}{3\kappa_C} \\ &= \frac{2\varepsilon_C}{3}. \end{aligned}$$

It remains to be shown that $\frac{2\varepsilon_C}{3} < f_j - \langle y_m, a_j\rangle$ for all $j \in \tilde{C}$.

Suppose there exists a $j \in \tilde{C}$ such that $f_j - \langle y_m, a_j\rangle \leq \frac{2\varepsilon_C}{3}$. Then

$$\begin{aligned} \mathrm{d}(P_h(y_0), b_j) &\leq \mathrm{d}(P_h(y_0), P_{b_j}(y_m)) \\ &\leq \mathrm{d}(P_h(y_0), y_m) + \mathrm{d}(y_m, P_{b_j}(y_m)) \\ &= \|y_m - P_h(y_0)\| + \mathrm{d}(y_m, b_j) \\ &< \frac{\varepsilon_C}{3\kappa_C} + \mathrm{d}(y_m, b_j) \\ &\leq \frac{\varepsilon_C}{3} + \mathrm{d}(y_m, b_j) \\ &= \frac{\varepsilon_C}{3} + |\langle y_m, a_j\rangle - f_j| \\ &\leq \frac{\varepsilon_C}{3} + \frac{2\varepsilon_C}{3} \\ &= \varepsilon_C. \end{aligned}$$

Thus there exists an element of b_j that is within ε_C of $P_h(y_0)$. This is a contradiction since $\mathcal{B}[P_h(y_0), \varepsilon_C] \subset \operatorname{int}(h_j)$. Therefore, for all $m > M$ and $j \in \tilde{C}$ we have that $2\|A_C^+\|\,\|r_m\| < f_j - \langle y_m, a_j \rangle$. ∎

THEOREM 2.16. *For all $m > M$, C satisfies properties* (2.10a)–(2.10c).

Proof. This is a consequence of Lemmas 2.11, 2.12, and 2.15. ∎

For the remainder of the section, we will assume that $C_m = C$ whenever $m > M$.

THEOREM 2.17. $\lim_{m\to\infty} \rho_m = 0$.

Proof. Let $m > M$. By assumption we have that $C_m = C$ and $\rho_m \le 2\kappa_C\|y_m - P_{b_C}(y_m)\|$. By Lemma 2.14, for all $m > M$, $P_{b_C}(y_m) = P_h(y_0)$, and thus $\rho_m \le 2\kappa_C\|y_m - P_h(y_0)\|$. Therefore

$$\lim_{m\to\infty} \rho_m \le \lim_{m\to\infty} 2\kappa_C\|y_m - P_h(y_0)\| = 0,$$

and $\lim_{m\to\infty} \rho_m = 0$. ∎

We have now established that it is possible to construct a sequence $\{\rho_m\}$ such that $\|y_m - P_h(y_0)\| \le \rho_m$ with $\lim_{m\to\infty} \rho_m = 0$. Next we exhibit a $y \in h$ such that $\|y - P_h(y_0)\| < \varepsilon$ for any $\varepsilon > 0$.

Given $\varepsilon > 0$, it is possible to determine an m such that C_m is nonempty and $\|y_m - P_h(y_0)\| < 2\varepsilon$. Bramley and Sameh have shown that using the conjugate-gradient method to solve $A_{C_m}^{\mathrm{T}} A_{C_m} y = A_{C_m}^{\mathrm{T}} b_{C_m}$ with an initial approximation of y_m results in a solution of $P_{b_{C_m}}(y_m)$ [2]. Using arguments in Lemma 2.9 and Theorem 2.10,

$$\|P_{b_{C_m}}(y_m) - P_h(y_0)\| < \varepsilon.$$

The following lemma shows that $P_{b_{C_m}}(y_m) \in h$.

LEMMA 2.18. *If $C_m \ne \emptyset$, then $P_{b_{C_m}}(y_m) \in h$.*

Proof. If $C_m = \{0, 1, \ldots, n-1\}$ (equivalently if $\tilde{C}_m = \emptyset$), then $P_b(y_m) \in b \subset h$. Next suppose that $\tilde{C}_m \ne \emptyset$. By Lemma 2.3, $P_{b_{C_m}}(y_m) \in \operatorname{int}(h_{\tilde{C}_m}) \subset h_{\tilde{C}_m}$. In addition $P_{b_{C_m}}(y_m) \in b_{C_m} \subset h_{C_m}$. Therefore $P_{b_{C_m}}(y_m) \in (h_{C_m} \cap h_{\tilde{C}_m}) = h$. ∎

Thus setting $y = P_{b_{C_m}}(y_m)$ specifies a point in h with the property that $\|y - P_h(y_0)\| < \varepsilon$. Moreover, if $m > M$, then by Lemma 2.14, $y = P_{b_{C_m}}(y_m)$ is in fact $P_h(y_0)$.

3. Implementation, applications, and numerical results. The two main results of section 2 were Theorem 2.10 and Theorem 2.17. In Theorem 2.10 we showed that if $C_m \ne \emptyset$, then $\|y_m - P_h(y_0)\| \le 2\|A_{C_m}^+\|\,\|r_m\|$, and we subsequently defined $\rho_m = 2\|A_{C_m}^+\|\,\|r_m\|$. Theorem 2.17 implied that for sufficiently large m, if $C_m = C$, then $\lim_{m\to\infty} \rho_m = 0$. We remark that for m sufficiently large and $C_m = C$, we have, using Lemma 2.14, $\|y_m - P_h(y_0)\| \le \frac{1}{2}\rho_m$.

Next we show how to construct a sequence $\{C_m\}$ so that for large enough m, $C_m = C$. Lemma 3.1 shows that if $C_m \ne \emptyset$, then $C \subset C_m$. As previously specified, $\mathcal{B}_m = \mathcal{B}[P_{b_{C_m}}(y_m), \mathrm{d}(y_m, P_{b_{C_m}}(y_m))]$, $C = \{i \mid P_h(y_0) \in b_i\}$, and C_m is selected to satisfy properties (2.10a)–(2.10c).

LEMMA 3.1. *If $C_m \ne \emptyset$, then $C \subset C_m$.*

Proof. The proof is by contradiction.

Suppose there exists an m such that $C - C_m \ne \emptyset$. Let $j \in C - C_m$ and hence $j \in \tilde{C}_m$. By Lemma 2.3 the closed ball $\mathcal{B}_m \subset \operatorname{int}(h_j)$. By Lemma 2.8, $\{y_{m+k}\}_{k=0}^{\infty} \subset \mathcal{B}_m$. Since $\mathcal{B}_m$ is closed and $\lim_{k\to\infty} y_{m+k} = P_h(y_0)$, $P_h(y_0) \in \mathcal{B}_m$. By definition of C, $P_h(y_0) \in b_j$. In addition we have shown above that $P_h(y_0) \in \mathcal{B}_m \subset \operatorname{int}(h_j)$. It is impossible for $P_h(y_0)$ to be both in the interior of h_j and on the boundary of h_j. Therefore, if $C_m \neq \emptyset$, we have that $C \subset C_m$. $\square$

By Theorem 2.16, $C_m = C$ satisfies properties (2.10a)–(2.10c) for sufficiently large m. By Lemma 3.1, if $C_m \neq \emptyset$, then $C \subset C_m$. Therefore, if C_m is chosen to have the smallest cardinality of any set satisfying properties (2.10a)–(2.10c), then for sufficiently large m, C_m must be the set C.

As previously stated, we compute

$$\rho_m = 2\|A^+_{C_m}\| \, \|r_m\|,$$

where

$$\|A^+_{C_m}\| = (\sigma_{\min}(A_{C_m}))^{-1},$$

$\sigma_{\min}(A_{C_m}) \neq 0$.

If the columns of A^{T} are linearly independent, we will subsequently show how to reduce the computational effort required to bound $\|y_m - P_h(y_0)\|$.

Proposition 3.2 may be directly inferred by the repeated application of the interlacing property for singular values [12].

PROPOSITION 3.2. *Let A^{T} have linearly independent columns and $S \subset \{0, 1, \ldots, n-1\}$, $S \neq \emptyset$; then $\|A^+_S\| \leq \|A^+\|$.*

We next determine a set C_m with the properties

$$f_{C_m} \in \mathrm{Range}(A_{C_m}), \tag{3.1a}$$

$$\text{for all } j \in \tilde{C}_m, \quad 2\|A^+\| \, \|r_m\| < f_j - \langle y_m, a_j \rangle, \tag{3.1b}$$

$$\begin{aligned} &\text{for all } j \in \tilde{C}_m, \quad e_{\pi_{m,j}} = \vec{0}, \\ &\text{where } \pi_{m,j} \in \{i \mid m-n \leq i \leq m-1 \ \text{ and } \ [i] = j\}. \end{aligned} \tag{3.1c}$$

C_m is well defined as $C_m = \{0, 1, \ldots, n-1\}$ vacuously satisfies properties (3.1a)–(3.1c). Using arguments similar to those presented in section 2, it can be shown that $\|y_m - P_h(y_0)\| \leq 2\|A^+\| \, \|r_m\|$. Thus, in order to bound the error of each Dykstra iterate we need only determine $\sigma_{\min}(A)$ and $\|r_m\|$.

PROPOSITION 3.3. *Let $S \subset \{0, 1, \ldots, n-1\}$, $S \neq \emptyset$, then $\|A_S\| \leq \|A\|$.*

Assuming that the columns of A^{T} are linearly independent, using Propositions 3.2 and 3.3 and selecting C_m to satisfy properties (3.1a)–(3.1c),

$$\begin{aligned} \|y_m - P_h(y_0)\| &\leq 2\|A^+\| \, \|r_m\| \\ &= 2\|A^+\| \, \|A_{C_m}(y_m - P_{h_{C_m}}(y_m))\| \\ &\leq 2\|A^+\| \, \|A_{C_m}\| \, \|y_m - P_{h_{C_m}}(y_m)\| \\ &\leq 2\|A^+\| \, \|A\| \, \|y_m - P_{h_{C_m}}(y_m)\| \\ &= 2\kappa\|y_m - P_{h_{C_m}}(y_m)\|, \end{aligned}$$

where $\kappa = \|A^+\| \, \|A\|$.

In order for the bound $2\|A^+\| \, \|r_m\|$ to be useful, $\{C_m\}$ must be selected such that $\lim_{m\to\infty} \|A^+\| \, \|r_m\| = 0$. To guarantee that

$$\lim_{m\to\infty} 2\kappa\|y_m - P_{h_{C_m}}(y_m)\| = 0,$$

it suffices that for sufficiently large m, $C_m = C$. As in section 2, with such a selection of $\{C_m\}$, $\lim_{m\to\infty} 2\|A^+\|\,\|r_m\| = 0$.

An important class of problems in statistical inference is to find the N-convex regression for a real valued function g_0 defined on a finite subset of $\Re$. Let $t_1 < t_2 < \cdots < t_q$, $q \geq N+1$. For a given $y_0 = (g_0(t_1), \ldots, g_0(t_q))^{\mathrm{T}}$ in $\Re^q$ we want to determine the best approximation to y_0 from the set of N-convex functions in $\Re^q$.

The real valued function g defined on $t_1, \ldots, t_q$ is N-convex if for any set of $N+1$ points $t_{i_1} < t_{i_2} < \cdots < t_{i_{N+1}}$, the Nth order divided difference, $g[t_{i_1}, \ldots, t_{i_{N+1}}] \geq 0$. This type of approximation problem generates a rectangular system of linear inequalities $Ay \leq f$ such that $\{y \mid Ay \leq f\}$ is a closed convex cone.

For N-convex regression problems it is possible to show that if $h = \{y \in \Re^q \mid y$ is N-convex$\}$, then $h = \cap_{i=0}^{q-N-1} h_i$, where $h_i = \{y \in \Re^q \mid \langle y, z_i\rangle \leq 0\}$, $b_i = \{y \in \Re^q \mid \langle y, z_i\rangle = 0\}$, and

$$\hat{z}_i(j) = \begin{cases} (-1)^{j-i+N}\binom{N}{j-i-1} & \text{whenever } \ 0 \leq j-i-1 \leq N, \\ 0 & \text{otherwise} \end{cases}$$

with

$$z_i = \frac{\hat{z}_i}{\|\hat{z}_i\|}. \tag{3.2}$$

The following proposition is easily proved.

PROPOSITION 3.4. *The* $\{z_i\}_{i=0}^{q-N-1}$ *as defined in* (3.2) *is linearly independent.*

Thus it is possible, given $y_0 = (g_0(t_1), \ldots, g_0(t_q))^{\mathrm{T}}$, to use Dykstra's algorithm to approximate $P_h(y_0)$ and to bound the norm of the error for each Dykstra iterate, $\|y_k - P_h(y_0)\|$, using the previously developed theory.

The set h of monotonically increasing functions can be expressed using a system of linear inequalities $h = \{y \in \Re^q \mid Ay \leq \vec{0}\}$, where $A^{\mathrm{T}} \in \Re^{q\times q-1}$. Let $\hat{e} \in \Re^q$ and for $i = 1, \ldots, q$, $j = 1, \ldots, q$, define

$$\hat{e}_i(j) = \begin{cases} 0 & \text{whenever } \ i = j, \\ 1 & \text{otherwise.} \end{cases}$$

Then $A^{\mathrm{T}} = [a_0, \ldots, a_{q-2}]$, where $a_i = \frac{1}{\sqrt{2}}(\hat{e}_{i+2} - \hat{e}_{i+1})$ for $i = 0, \ldots, q-2$.

By Proposition 3.4, the columns of A^{T} are linearly independent. In order to bound the norm of the error of the Dykstra iterates when approximating $P_h(y_0)$, we need only to calculate $\|A^+\|$ and $\|r_m\|$. As previously stated, $\|A^+\| = (\sigma_{\min}(A))^{-1}$ and since $(\sigma_{\min}(A))^{-1} = \lambda_{\min}(AA^{\mathrm{T}})^{-\frac{1}{2}}$, we need to estimate the smallest eigenvalue of AA^{T}. The matrix AA^{T} is tridiag$(-0.5, 1, -0.5)$, and thus $\lambda_{\min}(AA^{\mathrm{T}})$ can be approximated by using the QR algorithm with Givens rotations.

We present an example of monotonic regression and exhibit error bounds on $\|y_m - P_h(y_0)\|$. Here $y_0 = (g_0(t_1), \ldots, g_0(t_q))^{\mathrm{T}}$, where $g_0(t_i) \geq g_0(t_j)$ whenever $1 \leq i < j \leq q$. It can be shown that $P_h(y_0) = (g(t_1), \ldots, (g(t_1))^{\mathrm{T}}$, where $g(t_i) = \frac{1}{q}\sum_{j=1}^{q} g_0(t_j)$ for $1 \leq i \leq q$. Thus it is possible to compare $\|y_m - P_h(y_0)\|$ to the error bounds calculated using the previously derived theory. In the example $q = 31$, $g_0(t_i) = 16 - i$ for $1 \leq i \leq 31$ and $C = \{0, 1, \ldots, 30\}$. Table 3.1 shows $\|y_m - P_h(y_0)\|$ and $\rho_m = 2\|A^+\|\,\|r_m\|$ for selected values of m.

In order to compute the error bounds for this illustration, we need to determine for each iterate a set C_m that satisfies properties (3.1a)–(3.1c). For $m = 1, \ldots,$ it

Table 3.1.

m	$\|y_m - P_h(y_0)\|$	$\rho_m = 2\|A^+\|\ \|r_m\|$
6.0 E +3	6.3504 E +0	1.2778 E +1
1.2 E +4	8.1138 E −1	1.6326 E +0
1.8 E +4	1.0367 E −1	2.0854 E −1
2.4 E +4	1.3245 E −2	2.6651 E −2
3.0 E +4	1.6923 E −3	3.4051 E −3
3.6 E +4	2.1622 E −4	4.3506 E −4
4.2 E +4	2.7625 E −5	5.5586 E −5
4.8 E +4	3.5296 E −6	7.1020 E −6
5.4 E +4	4.5096 E −7	9.0740 E −7
6.0 E +e	5.7617 E −8	1.1593 E −7

can be shown that $C_m = \{0, 1, \ldots, 30\}$ satisfies properties (3.1a)–(3.1c). In addition, it is impossible to construct a proper subset of C_m which also satisfies these same properties. Thus, to compute a bound on $\|y_m - P_h(y_0)\|$ which tends to zero, we need only to calculate $\|r_m\| = \|Ay_m\|$ and multiply by the constant $2\|A^+\|$. In general, in order to guarantee that $\lim_{m\to\infty} \rho_m = 0$, the sequence of sets C_m must be selected judiciously. As previously discussed in this section, selecting C_m to have the smallest cardinality of any set satisfying properties (2.10a)–(2.10c) guarantees that $\lim_{m\to\infty} \rho_m = 0$.

The previous discussion shows that it is possible to compute a bound in the general Hilbert space setting on $\|x_m - P_H(x_0)\|$ that tends to zero where x_n is the nth iterate of Dykstra's algorithm and H is the intersection of a finite number of closed half spaces. The convergence is studied via an equivalent problem in $\Re^k$. This approach is well suited for those occasions where H is a closed convex cone uniquely determined by intersecting half spaces. The analysis is of particular interest to the author when using Dykstra's algorithm to estimate attribute utility values in nonmetric trade-off experiments.

REFERENCES

[1] S. Agmon, *The relaxation method for linear inequalities*, Canad. J. Math., 6 (1954), pp. 382–392.

[2] R. Bramley and A. Sameh, *Row projection methods for large nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 168–193.

[3] J. P. Boyle and R. L. Dykstra, *A method for finding projections onto the intersection of convex sets in Hilbert spaces*, in Advances in Order Restricted Statistical Inference, Lecture Notes in Statist. 37, Springer-Verlag, Berlin, 1985, pp. 28–47.

[4] R. L. Dykstra, *An algorithm for restricted least squares regression*, J. Amer. Statist. Assoc., 78 (1993), pp. 837–842.

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

[6] F. Deutsch and H. Hundal, *The rate of convergence of Dykstra's cyclic projections algorithm: The polyhedral case*, Numer. Funct. Anal. Optim., 15 (1994), pp. 537–565.

[7] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., John Wiley and Sons, Chichester, UK, 1991.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[9] S. P. Han, *A successive projection method*, Math. Program., 40 (1988), pp. 1–14.

[10] C. Hildreth, *A quadratic programming procedure*, Naval Res. Logist. Quart., 4 (1957), pp. 36 (erratum), 79–85.

[11] A. N. Iusem and A. R. De Pierro, *On the convergence properties of Hildreth's quadratic programming algorithm*, Math. Program., 47 (1990), pp. 37–51.

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge,

UK, 1985.

[13] B. F. Mitchell, V. F. Dem'yanov, and V. N. Malozemov, *Finding the point of a polyhedron closest to the origin*, SIAM J. Control, 12 (1974), pp. 19–26.

[14] T. S. Motzkin and J. Schoenberg, *The relaxation method for linear inequalities*, Canad. J. Math., 6 (1954), pp. 393–404.

[15] S. Xu, *Successive approximate algorithm for best approximation from a polyhedron*, J. Approx. Theory, 93 (1998), pp. 415–433.