Approximations in the l_∞ Norm and the Generalized Inverse

Delia Flores de Chela
Departamento de Matemáticas
Facultad de Ciencias
Universidad Central de Venezuela
Caracas, Venezuela

Submitted by Richard A. Brualdi

ABSTRACT

We introduce the concept of a strict l_∞-metric projector, based in the definition of strict approximation, to prove that for each matrix A of order $m\times n$ with coefficients in the field R of real numbers there exists a set of operators $G: R^m \to R^n$ homogeneous and continuous, but not necessarily linear (strict generalized inverse) such that $AGA=A$ and $\|AGy-y\|$ is minimized for all y, when the norm is the l_∞ norm. We investigate the properties of these operators and prove that there are two distinguished operators $A_{\infty,\beta}^{-1}$ and A_∞^{-1} which are extensions of the generalized inverse introduced by Newman and Odell in the case of a strictly convex norm.

INTRODUCTION

Let $R^{m\times n}$ denote the set of matrices of order $m\times n$ with coefficients in R. Let $A\in R^{m\times n}$ and let $\|\cdot\|$ be a norm in R^m.

From the results of R. Penrose [7] and C. R. Rao and S. K. Mitra [8] it follows that when $\|\cdot\|$ is an Euclidean norm in R^m, there exists $G\in R^{n\times m}$ such that

(i) $AGA=A$,
(ii) $GAG=G$,
(iii) Gy is a $\|\cdot\|$-best approximation of $Ax=y$ for all y.

Moreover, if $\|\cdot\|$ is the usual Euclidean norm, there exists a unique matrix G, the Moore-Penrose inverse of A, which satisfies (i), (ii), (iii) and

(iv) Gy has minimum norm among the $\|\cdot\|$-best approximations of $Ax=y$, for all y.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/010003+19$02.75

But if the norm is the l_∞ norm

$$\|x\| = \max_i |x_i|,$$

it is known [4, Proposition 4] that there does not exist a matrix G satisfying (i), (ii), and (iii) for every matrix A.

On the other hand, the results of Newman and Odell [6] show that if the norm $\|\cdot\|$ is a strictly convex norm (i.e., $\|x+y\| = \|x\| + \|y\|$ with $x \neq 0$ holds if and only if $y = \mu x$ for some $\mu \geq 0$), then there exists an operator

$$G: R^m \to R^n,$$

homogeneous and continuous, but not necessarily linear, such that (i), (ii), and (iii) hold. All the l_p norms given by

$$\|x\| = \left(\sum |x_i|^p\right)^{1/p}, \qquad p > 1, \tag{1}$$

are strictly convex, but the l_∞ norm is not. So the following question arises: Does there exist an operator $G: R^m \to R^n$, homogeneous and continuous, but not necessarily additive, such that (i), (ii), and (iii) hold for the l_∞ norm?

It is the purpose of this paper to prove that the answer to this question is in the affirmative. We shall prove that there exists a set of such operators and that among them there is one that may be considered as an extension of the generalized inverse introduced by Newman and Odell in the case of a strictly convex norm.

For our purpose it is not restrictive to assume that

$$A = (B : BS) \tag{2}$$

where $B \in R^{m \times r}$ has rank r and S is a certain matrix. In fact, given $A \in R^{m \times n}$ of rank r, there exists a permutation matrix P such that $\hat{A} = AP$ has the form (2). If there exists

$$G: R^m \to R^n,$$

homogeneous and continuous, such that

(a) $\hat{A}G\hat{A} = \hat{A}$,
(b) $G\hat{A}G = G$,
(c) Gy is an l_∞-best approximation of $\hat{A}x = y$,

for all y, then the map

$$PG: R^m \to R^n$$

is homogeneous and continuous and satisfies (i), (ii), and (iii) when the norm is the l_∞ norm.

1. BASIC DEFINITIONS AND KNOWN RESULTS

Let $A \in R^{m \times n}$, where $m \geqslant n$, and suppose A is of rank r. Consider the system of linear equations

$$Ax = y. \qquad (\mathrm{I}_1)$$

Let $\|\cdot\|$ be a norm in R^m. A $\|\cdot\|$-best approximation of (I_1) is a vector x_0 in R^n which satisfies the relation

$$\|Ax_0 - y\| = \inf_x \|Ax - y\|.$$

Thus, Ax_0 is a $\|\cdot\|$-closest point to y in the range of A. Since the range of A is a closed set in R^m, a $\|\cdot\|$-best approximation of (I_1) always exists.

PROPOSITION 1. *Suppose R^n is endowed with the norm l_p, $p > 1$, given by* (1). *Then*:

(i) *If* rank $A = n$, (I_1) *has a unique l_p-best approximation.*

(ii) *If* rank $A = r$ *and A is given by* (2), *the set of l_p-best approximations of* (I_1) *is the linear manifold*

$$x_0 + \operatorname{Ker} A$$

where $x_0 = (x_1, \dots, x_r, 0, \dots, 0)^T$ *is the transpose of* $(x_1, \dots, x_r, 0, \dots, 0)$ *and* $(x_1, \dots, x_r)^T$ *is the unique l_p-best approximation of the system*

$$B\hat{x} = y.$$

Proof. Part (i) is a consequence of Lemma 2 in [1, p. 129]. Part (ii) follows from (i) and the definition of $\|\cdot\|$-best approximation. ■

Let $S=\{i_1,\dots,i_s\}$ be a subset of $E=\{1,2,\dots,m\}$. The system (I_1) is inconsistent over S if

$$\begin{pmatrix} A^{i_1} \\ \vdots \\ A^{i_s} \end{pmatrix} x = \begin{pmatrix} y_{i_1} \\ \vdots \\ y_{i_s} \end{pmatrix}$$

is inconsistent, where A^j represent the jth row of A.

Let

$$r(x)=Ax-y$$
$$=(r_1(x),\dots,r_m(x))^T.$$

Suppose M is a linear manifold in R^n. We say that b is an l_∞-best approximation or a Chebyshev approximation of $Ax=y$ on S with respect to M if

$$\inf_{x\in M}\Big(\max_{i\in S}|r_i(x)|\Big)=\max_{i\in S}|r_i(b)|.$$

The deviation of $Ax=y$ on S with respect to M is given by

$$\rho(S)=\inf_{x\in M}\Big(\max_{i\in S}|r_i(x)|\Big).$$

When M is fixed, we call $\rho(S)$ and b simply the deviation of $Ax=y$ on S and the best approximation of $Ax=y$ on S, respectively.

If $W(S)$ is the set of all best approximations of $AX=y$ on S, we define the characteristic set of $Ax=y$ on S as the set

$$C(S)=\{i\in S:|r_i(b)|=\rho(S)\text{ for all } b\in W(S)\}.$$

Proposition 2 [3, Lemma 1; 2, Theorems 4, 5, 7].

(i) *The set* $W(S)$ *is nonempty.*
(ii) *The set* $C(S)$ *is nonempty.*

The next proposition follows from the properties of $W(S)$ and of convex sets.

PROPOSITION 3. *The set $W(S)$ is convex.*

COROLLARY 1. *Suppose the system $Ax=y$ is inconsistent over S. Then if $i\in C(S)$,*

$$\operatorname{sgn}(r_i(b))=\operatorname{sgn}(r_i(b_0))$$

for all b, b_0 in $W(S)$.

Proof. Suppose that $i\in C(S)$ and that there are $b, b_0 \in W(S)$ such that

$$\operatorname{sgn}(r_i(b))\neq\operatorname{sgn}(r_i(b_0));$$

then $r_i(b)=-r_i(b_0)$ and

$$r_i\left(\tfrac{1}{2}b+\tfrac{1}{2}b_0\right)=\tfrac{1}{2}\left[r_i(b)+r_i(b_0)\right]=0.$$

Now, that $W(S)$ is convex implies that $\frac{1}{2}b+\frac{1}{2}b_0 \in W(S)$. Therefore, since $Ax=y$ is inconsistent over S, we have that

$$\left|r_i\left(\tfrac{1}{2}b+\tfrac{1}{2}b_0\right)\right|=\rho(S)>0,$$

which contradicts the relation $r_i(\frac{1}{2}b+\frac{1}{2}b_0)=0$ and proves the corollary. ■

The concept of *strict approximation*, due to J. R. Rice, has been formulated in a constructive manner by him [9] and by J. Descloux [3]. They consider the approximation of a real function by a set of given real functions $f_1,\dots,f_n$ defined on a finite set. We shall follow their ideas to construct the strict approximation of an inconsistent system of the form (I_1).

Step 1. Find the l_∞-best approximations of (I_1) on E with respect to R^n. Let ρ_1, W_1, and C_1 be the deviation, the set of l_∞-best approximations, and the characteristic set, respectively, of (I_1) on E. Then

$$\rho_1=\inf_{x\in R^n}(\max|r_i(x)|) \qquad (i\in E),$$

and the vectors of W_1 satisfy

$$M_1: r_\alpha(X)=\varepsilon(\alpha)\rho_1, \qquad \alpha\in C_1, \tag{a_1}$$

where, by Corollary 1, either $\varepsilon(\alpha)=1$, or else $\varepsilon(\alpha)=-1$ for all $x\in W_1$.

Suppose

$$C_1 = \{\alpha_1, \ldots, \alpha_t\}.$$

Let

$$^{\langle 1 \rangle}A = \begin{pmatrix} A^{\alpha_1} \\ \vdots \\ A^{\alpha_t} \end{pmatrix}.$$

Then, the set of solutions of (a_1) is a linear manifold

$$M_1 = x_0 + \mathrm{Ker}\,^{\langle 1 \rangle}A,$$

where $x_0 \in W_1$. If $s_1 = \mathrm{rank}\,^{\langle 1 \rangle}A$, then $\dim M_1 = n - s_1$.

Step 2. Find the l_∞-best approximations of (I_1) on $E \setminus C_1$ with respect to M_1. We associate with this problem the real number

$$\rho_2 = \inf_{x \in M_1} (\max |r_i(x)|) \qquad (i \in E \setminus C_1),$$

the set W_2 of l_∞-best approximations, and the characteristic set C_2 of (I_1) on $E \setminus C_1$, with respect to M_1.

We define the manifold M_2 as the set of all vectors satisfying

$$M_2 : \begin{cases} r_\alpha(x) = \varepsilon(\alpha)\rho_1, & \alpha \in C_1, \quad (a_1) \\ r_\beta(x) = \varepsilon(\beta)\rho_2, & \beta \in C_2. \quad (a_2) \end{cases}$$

Note that $C_2 \subseteq E \setminus C_1$, so $C_1 \cap C_2 = \varnothing$. If $C_2 = \{\beta_1, \ldots, \beta_q\}$ and

$$^{\langle 2 \rangle}A = \begin{bmatrix} A^{\alpha_1} \\ \vdots \\ A^{\alpha_t} \\ A^{\beta_1} \\ \vdots \\ A^{\beta_q} \end{bmatrix},$$

then the solution manifold of the system is

$$M_2 = x_1 + \mathrm{Ker}\,^{\langle 2 \rangle}A,$$

where $x_1 \in W_2$. Suppose that the rank of ${}^{\langle 2\rangle}A$ is equal to s_1+s_2, so that

$$\dim M_2 = n - s_1 - s_2.$$

If $\rho_2 = 0$, then $r_j(x) = 0$, where j runs over $E \backslash C_1$, by the definition of ρ_2. In that case $C_2 = E\backslash C_1$. All the m rows of A are involved in the definition of M_2. Therefore the system is of rank n, and in consequence has a unique solution. If $\rho_2 \neq 0$, $n - s_1 - s_2 > 0$, and $C_2 \neq E\backslash C_1$, the process is continued until we reach the Kth step. Find the l_∞-best approximations of (I_1) on $E\backslash C_1 \backslash \cdots \backslash C_{K-1}$ with respect to M_{K-1}.

Define ρ_K, W_K, C_K, and the manifold M_K determined by the set of solutions of

$$M_K: \begin{cases} r_\alpha(x) = \varepsilon(\alpha)\rho_1, & \alpha \in C_1, \qquad (\mathrm{a}_1) \\ r_\beta(x) = \varepsilon(\beta)\rho_2, & \beta \in C_2, \qquad (\mathrm{a}_2) \\ \quad\vdots & \\ r_\tau(x) = \varepsilon(\tau)\rho_K, & \tau \in C_K, \qquad (\mathrm{a}_K) \end{cases}$$

where $\dim M_K = n - s_1 - s_2 - \cdots - s_K = 0$. This must happen eventually, because we are using more rows of A at each step and the m rows are of rank n. The manifold M_K consists of precisely one vector x^*, which is the common solution of the systems (a_1), $(\mathrm{a}_2), \ldots, (\mathrm{a}_K)$.

We define x^* to be the *strict approximation* of $Ax = y$.

REMARK 1. The strict approximation is the only best approximation of $Ax = y$ that is also a best approximation of each of the subsystems

$$\langle A^i, x\rangle = y_i, \qquad i \in E\backslash C_1 \backslash \cdots \backslash C_j, \quad j = 1, \ldots, K.$$

The deviations $\rho_1, \ldots, \rho_K$ of this system satisfy the relation $\rho_1 > \cdots > \rho_K$ (for proof refer to [3]).

When rank $A = r$ and A is given by (2), we define the strict approximations of A as the linear manifold $x^* + \ker A$, where $x^* = (x_1, \ldots, x_r, 0, \ldots, 0)^T$, and $(x_1, \ldots, x_r)^T$ is the unique strict approximation of the system

$$B\hat{x} = y.$$

THEOREM 1. *Let $A \in R^{m\times n}$ be of* rank n. *For each $p = 2, 3, \ldots,$ let x_p be the unique l_p-best approximation and let x^* be the strict approximation of the* *inconsistent system* $Ax=y$. *Then*

$$\lim_{p\to\infty} x_p = x^*.$$

For proof refer to [3].

COROLLARY 2. *Let* $A\in R^{m\times n}$, *let* x_p *be any* l_p*-best approximation of a system of equations of the form* (I_1), *and let* x^* *be any strict approximation of* (I_1). *Then*

$$\lim_{p\to\infty} Ax_p = Ax^*.$$

Proof. If $\operatorname{rank} A=n$, the corollary follows directly from Theorem 1. Suppose that $\operatorname{rank} A=r$ and $A=(B\colon BS)$, where $B\in R^{m\times r}$ has rank r. Then if

$$\tilde{x}_p=(x_{1,p},\dots,x_{r,p})^T \quad\text{and}\quad \tilde{x}^*=(x_1,\dots,x_r)^T$$

are the unique l_p-best approximation and the strict approximation of

$$Bx=y$$

respectively, then

$$\hat{x}_p=(x_{1,p},\dots,x_{r,p},0,\dots,0)^T$$

and

$$\hat{x}^*=(x_1,\dots,x_r,0,0,\dots,0)^T$$

are an l_p-best approximation and a strict approximation of (I_1), respectively. Moreover, the sets of l_p-best approximations and strict approximations of (I_1) are, respectively

$$\hat{x}_p+\operatorname{Ker} A \tag{3}$$

and

$$\hat{x}^*+\operatorname{Ker} A. \tag{4}$$

Now, by Theorem 1,

$$\lim_{p\to\infty} \tilde{x}_p = \tilde{x}^*.$$

Therefore

$$\lim_{p\to\infty} \hat{x}_p = \hat{x}^*,$$

which implies

$$\lim_{p\to\infty} A\hat{x}_p = A\hat{x}^*.$$

This relation, together with (3) and (4), implies our corollary. ■

2. THE STRICT l_∞-METRIC PROJECTOR

Let L be a subspace of R^m, and ϕ a norm in R^m. The ϕ-metric projector on L is the point to set-valued mapping

$$P_{L,\phi}R^m \to R^m$$

whose image is contained in L and which associates with each y in R^m the set of ϕ-closest points to y in L. Thus

$$P_{L,\phi}(y) = \left\{x_0 \in L : \phi(y - x_0) = \inf_{x\in L} \phi(y-x)\right\}$$

[1, p. 157].

When ϕ is a strictly convex norm, $P_{L,\phi}$ is an ordinary function which assigns to each $y \in R^m$ its unique ϕ-closest point in L.

THEOREM 2. *If ϕ is a strictly convex norm, then for any subspace L in R^m and any point y in R^m we have that*

(a) $P_{L,\phi}(y) = y$ *if and only if* $y \in L$,
(b) $P^2_{L,\phi} = P_{L,\phi}$,
(c) $P_{L,\phi}(\lambda y) = \lambda P_{L,\phi}(y)$,
(d) $P_{L,\phi}(x+y) = x + P_{L,\phi}(y)$ *for all* $x \in L$,
(e) $P_{L,\phi}$ *is continuous on* R^m.

For proof refer to [1, p. 130].

PROPOSITION 4. *If* $\phi = l_p$, $p>2$, *then* $P_{L,\phi}$ *is uniformly continuous on* R^m.

For proof refer to [5, p. 236].

PROPOSITION 5. *If L is a subspace of* R^m *of dimension r, then there exists a matrix* $A_L \in R^{m\times r}$ *of* rank *r such that* $y \in R^m$ *belongs to L if and only if the system*

$$A_L x = y \tag{I_2}$$

is consistent.

Proof. Let $l_1, \dots, l_r$ be a basis for L. If $e_1, e_2, \dots, e_m$ is the canonical basis of R^m,

$$l_i = \sum_{j=1}^{m} a_{ji} e_j, \qquad i = 1, \dots, r.$$

Let A_L be

$$A_L = (a_{ji});$$

then $A_L \in R^{m\times r}$, rank $A_L = r$, and it is easy to prove that (I_2) is consistent if and only if $y \in L$. ∎

The matrix A_L is called a parametric representation of L.

REMARK 2. If A_L and B_L are two parametric representations of L, then

$$B_L = A_L \cdot P,$$

where P is a nonsingular matrix of order r.

REMARK 3. Let ϕ be the l_∞ norm in R^m, and let L be a subspace of R^m. If $y \in R^m \setminus L$, then $A_L x^*$ is in $P_{L,\phi}(y)$, where x^* is the unique strict approximation of the inconsistent system $A_L x = y$.

REMARK 4. Let $A \in R^{m\times r}$, and let P be a nonsingular matrix of order r. Then x is an l_∞-best approximation of $Ax = y$ if and only if $P^{-1}x$ is an l_∞-best approximation of $APx=y$. Hence, by the definition of strict approximation, x^* is the strict approximation of $Ax=y$ if and only if $P^{-1}x^*$ is the strict approximation of $APx=y$.

Assuming that R^m is endowed with the l_∞ norm, we use Remarks 2, 3, and 4 to define for each subspace L of R^m a function

$$\bar{P}_L: R^m \to R^m$$

by

$$\bar{P}_L(y)=\begin{cases} y, & y\in L \\ A_L x^*, & y\notin L. \end{cases}$$

It should be noted that the image of $\bar{P}_L$ lies in L. We call $\bar{P}_L$ the strict l_∞-metric projector.

In the following propositions we shall develop some properties of $\bar{P}_L$.

PROPOSITION 6. *If $R(A)$ is the range of A, then*

(i) $\bar{P}_{R(A)}A=A$.

(ii) *If $y\notin R(A)$,*

$$P_{R(A)}(y)=Ax^*$$

for any strict approximation x^* of $Ax=y$.

(iii) $AG\bar{P}_{R(A)}=\bar{P}_{R(A)}$ *for all generalized inverses G of A.*

Proof. (i): By the definition of $\bar{P}_{R(A)}$ we have that $\bar{P}_{R(A)}A=A$.

(ii): If rank $A=n$, we may suppose that $A_{R(A)}=A$. Therefore, if $y\notin R(A)$ we have that $\bar{P}_{R(A)}(y)=Ax^*$, where x^* is the unique strict approximation of $Ax=y$. Next suppose that rank $A<n$ and that A is given by (2), that is

$$A=(B:BS)$$

where $B\in R^{m\times r}$ has rank r. Then we may suppose that $A_{R(A)}=B$.

Therefore, if $y\notin R(A)$, we have that $\bar{P}_{R(A)}(y)=B\bar{x}^*$ where $\bar{x}^*=(x_1,\ldots,x_r)^T$ is the unique strict approximation of $B\hat{x}=y$. It is easy to show that if

$$\hat{x}^*=(x_1,\ldots,x_r,0,\ldots,0)^T,$$

then

$$B\bar{x}^* = A\hat{x}^* = Ax^*$$

for any strict approximation x^* of $Ax = y$. Hence, if $y \notin R(A)$,

$$\bar{P}_{R(A)}(y) = Ax^*$$

for any strict approximation x^* of $Ax = y$, which proves (ii).

(iii): If $y \in R(A)$,

$$\begin{aligned} AG\bar{P}_{R(A)}(y) &= AGAx^* \\ &= Ax^* \\ &= P_{R(A)}(y). \end{aligned}$$

If $y \in R(A)$,

$$\bar{P}_{R(A)}(y) = Aw$$

for some $w \in R^n$. Hence,

$$\begin{aligned} AG\bar{P}_{R(A)}(y) &= AGAw = Aw \\ &= \bar{P}_{R(A)}(y). \end{aligned}$$

Therefore, for all generalized inverses G of A, we have

$$AG\bar{P}_{R(A)} = \bar{P}_{R(A)}. \qquad \blacksquare$$

PROPOSITION 7. *If $P_{L,p}$ is the metric projector associated with a subspace L in R^m and the l_p norm, $p > 1$, in R^m, then*

$$\lim_{p \to \infty} P_{L,p} = \bar{P}_L$$

pointwise.

Proof. Let $y \in R^m \setminus L$. Let y_p be the l_p-closest point to y in L, and x_p be the l_p-best approximation of the inconsistent system

$$A_L x = y.$$

Then

$$P_{L,p}(y) = A_L x_p = y_p.$$

Hence,

$$\lim_{p\to\infty} P_{L,p}(y) = \lim_{p\to\infty} A_L x_p$$
$$= A_L x^*,$$

by Theorem 1. Thus

$$\lim_{p\to\infty} P_{L,p}(y) = \bar{P}_L(y), \qquad y \notin L.$$

If $y \in L$, then

$$P_{L,p}(y) = \bar{P}_L(y) = y \qquad \text{for all } p.$$

Therefore

$$\lim_{p\to\infty} P_{L,p}(y) = \bar{P}_L(y) \qquad \text{for all } y,$$

which proves the proposition. ■

The following result is a consequence of Theorem 2 and Proposition 7.

PROPOSITION 8. *For any subspace L of R^m and $y \in R^m$ we have that:*

(a) $\bar{P}_L(y) = y$ *if and only if* $y \in L$.
(b) $\bar{P}_L^2 = \bar{P}_L$.
(c) $\bar{P}_L(\lambda y) = \lambda \bar{P}_L(y)$ *for all* $\lambda \in R$.
(d) $\bar{P}_L(x+y) = x + \bar{P}_L(y)$ *for all* $x \in L$.

PROPOSITION 9. *For any subspace L of R^m, $\bar{P}_L$ is uniformly continuous.*

Proof. We have that $\bar{P}_L = \lim_{p\to\infty} P_{L,p}$ pointwise. Hence, given $\varepsilon > 0$ and $y_1, y_2 \in R^m$, there exists N such that $p > N$ implies that

$$\left\|\bar{P}_L(y_1) - P_{L,p}(y_1)\right\|_\infty < \frac{\varepsilon}{3} \quad \text{and} \quad \left\|\bar{P}_L(y_2) - P_{L,p}(y_2)\right\|_\infty < \frac{\varepsilon}{3}. \tag{5}$$

Let $p_0 > N$. Since P_{L,p_0} is uniformly continuous, by Proposition 4, there exists $\delta > 0$ such that

$$\|y_1 - y_2\|_\infty < \delta \quad \Rightarrow \quad \|P_{L,p_0}(y_1) - P_{L,p_0}(y_2)\|_\infty < \varepsilon/3. \tag{6}$$

Now,

$$\begin{aligned} \|\bar{P}_L(y_1) - \bar{P}_L(y_2)\|_\infty &\leq \|\bar{P}_L(y_1) - P_{L,p_0}(y_1)\|_\infty \\ &\quad + \|P_{L,p_0}(y_1) - P_{L,p_0}(y_2)\|_\infty \\ &\quad + \|P_{L,p_0}(y_2) - \bar{P}_L(y_2)\|_\infty. \end{aligned}$$

This last relation together with (5) and (6) implies that, given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|y_1 - y_2\|_\infty < \delta \quad \Rightarrow \quad \|\bar{P}_L(y_1) - \bar{P}_L(y_2)\| < \varepsilon.$$

Thus, $\bar{P}_L$ is uniformly continuous. ∎

3. THE STRICT GENERALIZED INVERSE

Let $A \in R^{m\times n}$. A strict generalized inverse (s.g.i.) of A is a homogeneous and continuous operator X from R^m to R^n such that:

(i) $AXA = A$.
(ii) $XAX = X$.
(iii) $X(y)$ is a strict approximation of a system of equations of the form (I_1), for every $y \notin R(A)$.

PROPOSITION 10. *Every s.g.i. X of A satisfies the equation*

$$AX = \bar{P}_{R(A)}.$$

Proof. If $y \notin R(A)$, then $X(y)$ is a strict approximation of (I_1), by (iii). Therefore, according to the definition of $\bar{P}_{R(A)}$,

$$AX(y) = \bar{P}_{R(A)}(y). \tag{7}$$

If $y \in R(A)$, then $y = Az$ for some $z \in R^n$. Hence

$$\bar{P}_{R(A)}(y) = y = Az$$

and

$$\begin{aligned} AX(y) &= AXAz \\ &= Az, \end{aligned}$$

by (i). So

$$AX(y) = \bar{P}_{R(A)}(y).$$

Combining this with (7), we obtain that

$$AX(y) = \bar{P}_{R(A)}(y)$$

for all $y \in R^m$. Thus,

$$AX = \bar{P}_{R(A)}.$$

THEOREM 3. *Let S be the set of homogeneous and continuous operators G from R^m to R^n such that $AGA = A$. Then*

$$\left\{ G\bar{P}_{R(A)} \textit{ for all } G \in S \right\}$$

is the set of strict generalized inverses of A.

Proof. Let $G \in S$ and

$$X = G\bar{P}_{R(A)}.$$

It is clear that

$$X: R^m \to R^n$$

is homogeneous and continuous. Next,

$$\begin{aligned} AXA &= A\left(G\bar{P}_{R(A)}\right)A \\ &= AG\left(\bar{P}_{R(A)}A\right) \\ &= AGA, \end{aligned}$$

by (i) of Proposition 6, which implies that $AXA = A$. And

$$\begin{aligned} XAX &= \left(G\bar{P}_{R(A)}\right)A\left(G\bar{P}_{R(A)}\right) \\ &= G\left(\bar{P}_{R(A)}A\right)G\bar{P}_{R(A)} \\ &= GAG\bar{P}_{R(A)} \\ &= G\left(AG\bar{P}_{R(A)}\right) \\ &= G\bar{P}_{R(A)}, \end{aligned}$$

by (iii) of Proposition 6. Hence,

$$XAX = X.$$

Finally, if $y \notin R(A)$,

$$\begin{aligned} X(y) &= G\bar{P}_{R(A)}(y) \\ &= GAx^* \end{aligned}$$

for some strict approximation x^* of $Ax = y$. Hence,

$$\begin{aligned} AX(y) &= AGAx^* \\ &= Ax^*. \end{aligned}$$

Thus,

$$X(y)=x^*+z,$$

$z\in\operatorname{Ker}A$. Hence $X(y)$ is a strict approximation of $Ax=y$ for every $y\notin R(A)$.

Conversely, if X is a s.g.i. of A, then $AX=\bar{P}_{R(A)}$. Therefore $XAX=X\bar{P}_{R(A)}$, which implies, by (ii) of the definition of s.g.i., that $X=X\bar{P}_{R(A)}$. ■

Corollary 4. *If* rank $A=n$, *there exists a unique strict generalized inverse* A_∞^+ *of* A, *given by*

$$A_\infty^+=A^+\bar{P}_{R(A)},$$

where A^+ *is the Moore-Penrose inverse of* A.

Proof. From Theorem 3 it is clear that

$$A_\infty^+=A^+\bar{P}_{R(A)}$$

is a s.g.i. of A.

Now, if X and Y are s.g.i. of A, then

$$AX=AY=\bar{P}_{R(A)},$$

by Proposition 10. But since A has rank n, the last relation implies

$$X=Y,$$

which proves our corollary. ■

4. A DISTINGUISHED STRICT GENERALIZED INVERSE

Suppose that $A\in R^{m\times n}$ has rank less than n. Then the strict approximation of $Ax=y$ is not unique. But if we consider a strictly convex norm β or the l_∞ norm in R^n, we may distinguish a strict approximation of

$$Ax=y. \qquad (I_1)$$

In fact, the linear manifold

$$H=\left\{A^{+}\bar{P}_{R(A)}(y)+z : z\in \operatorname{Ker} A\right\}$$

is the set of strict approximations of (I_1). Moreover, since

$$P_{H,\beta}(0)=P_{\operatorname{Ker} A,\beta}\left(-A^{+}\bar{P}_{R(A)}(y)\right)+A^{+}\bar{P}_{R(A)}(y),$$

we have that

$$A^{+}\bar{P}_{R(A)}(y)+P_{\operatorname{Ker} A,\beta}\left(-A^{+}\bar{P}_{R(A)}(y)\right)=\left(I-P_{\operatorname{Ker} A,\beta}\right)A^{+}\bar{P}_{R(A)}(y)$$

is the unique element of H of minimum β-norm.

Similarly, by the definition of $\bar{P}_{\operatorname{Ker} A}$ we have that

$$A^{+}\bar{P}_{R(A)}(y)+\bar{P}_{\operatorname{Ker} A}\left(-A^{+}\bar{P}_{R(A)}(y)\right)=\left(I-\bar{P}_{\operatorname{Ker} A}\right)A^{+}\bar{P}_{R(A)}(y)$$

is a strict approximation of (I_1) of minimum l_∞-norm.

THEOREM 4. *The operators*

$$A_{\beta,\infty}^{-1}=\left(I-P_{\operatorname{Ker} A,\beta}\right)A^{+}\bar{P}_{R(A)}$$

and

$$A_{\infty}^{-1}=\left(I-\bar{P}_{\operatorname{Ker} A}\right)A^{+}\bar{P}_{R(A)}$$

are s.g.i. of A *such that if* y *is not in the range of* A, *then* $A_{\beta,\infty}^{-1}(y)$ *is the unique strict approximation of* (I_1) *of minimum* β*-norm and* $A_{\infty}^{-1}(y)$ *is a strict approximation of* (I_1) *of minimum* l_∞*-norm. When* y *is in the range of* A, $A_{\beta,\infty}^{-1}(y)$ *is the unique solution of* (I_1) *of minimum* β*-norm and* $A_{\infty}^{-1}(y)$ *is a solution of* (I_1) *of minimum* l_∞*-norm.*

Proof. This follows immediately from the definitions of $A_{\beta,\infty}^{-1}$ and A_{∞}^{-1}. ∎

REMARK 5. Since, by Remark 1, the strict approximations of $Ax=y$ may be considered as the best among its l_∞-best approximations, the result of

Theorem 4 says that $A_{\beta,\infty}^{-1}$ and A_∞^{-1} are the natural extensions of the operator $A_{\beta,\alpha}^{-1}$, p-q generalized inverse of Newman and Odell [6], when $A\in R^{m\times n}$, $\alpha=l_\infty$ and β is an essentially strictly convex norm or the l_∞ norm.

Most of the results of this paper are part of a Ph.D. dissertation submitted in July 1978 to the University of Sussex, Falmer, England. I thank Professor Walter Ledermann, my supervisor, for his guidance and encouragement.

REFERENCES

1 A. Ben-Israel and N. E. Greville, *Generalized Inverses. Theory and Applications*, Wiley-Interscience, New York, 1973.
2 J. Descloux, Contribution au calcul des approximations de Tschebicheff, Doctoral Dissertation 15–18, Lousanne, 1961.
3 J. Descloux, Approximation in L^p and Chebyshev approximations, *SIAM J. Appl. Math.* 11:1017–1026 (1963).
4 D. Flores de Chela, Generalized inverses on normed vector spaces, *Linear Algebra and Appl.* 26:243–263 (1979).
5 R. B. Holmes and B. Kripke, Smoothness of approximation, *Michigan Math. J.* 15:225–248 (1966).
6 G. Newman and P. Odell, On the concept of p-q generalized inverse, *SIAM J. Appl. Math.* 17:520–525 (1969).
7 R. Penrose, On best approximate solution of linear matrix equations, *Proc. Cambridge Philos. Soc.* 52:17–19 (1955).
8 C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.
9 J. Rice, Tchebyckeff approximation in a compact metric space, *Bull. Amer. Math. Soc.* 68:405–410 (1968).

Received 19 November 1980; revised 9 March 1981