JOURNAL OF APPROXIMATION THEORY **88**, 209–227 (1997)
ARTICLE NO. AT963024

Orthant–Monotonic Norms and Overdetermined Linear Systems

K. Ziętak

Institute of Computer Science, University of Wrocław, ul. Przesmyckiego 20, 51-151 Wrocław, Poland
E-mail: zietak@ii.uni.wroc.pl

Communicated by E. W. Cheney

Received January 23, 1995; accepted February 27, 1996

In the paper the properties of $\|\cdot\|$-approximate solutions of real overdetermined linear systems $Ax=b$ are investigated. We characterize the norms for which the approximate solutions are in the convex hull of the basic solutions of $Ax=b$. For this purpose we prove some properties of the orthant–monotonic norms. We derive an explicit formula which expresses the approximate solutions of $Ax=b$, with respect to some norms, as convex combinations of the basic solutions for an $n+1$ by n real matrix A of rank n. Moreover, we consider the relation between the sets of the weighted approximate solutions for the orthant–monotonic norms satisfying some additional conditions.

1. INTRODUCTION

A series of papers has recently appeared which has taken a new look at the geometric properties of approximate solutions of overdetermined linear systems (see for example Ben-Israel [2], Ben-Tal and Teboulle [3], Berg [4], and Miao and Ben-Israel [15, 16]). In this paper we generalize some of the results presented in these papers.

Let $\mathscr{R}_r^{m\times n}$ denote the set of real $m\times n$ matrices A of rank r. If A is of arbitrary rank then we write $A\in\mathscr{R}^{m\times n}$. We consider the linear system $Ax=b$, $A\in\mathscr{R}^{m\times n}$ and $b\in\mathscr{R}^m$. If this system is inconsistent then we find its approximate solution x with respect to a norm $\|\cdot\|$. Thus we deal with the problem

$$\min_{x\in\mathscr{R}^n}\|Ax-b\|. \tag{1}$$

We do not assume that a norm is strictly homogeneous. Therefore we have

$$\|x\|\geqslant 0 \quad \text{with equality iff } x=0, \qquad \|x+y\|\leqslant\|x\|+\|y\|,$$

0021-9045/97 $25.00

and

$$\|\lambda x\| = \lambda \, \|x\| \qquad \text{for all real positive } \lambda. \tag{2}$$

The minimum (1) is denoted by δ. We assume that b does not belong to the subspace spanned by columns of A. Then $\delta > 0$.

A solution of (1) is often called an $\|\cdot\|$-*approximate solution.* In particular, the l_2-approximate solutions are the *least squares solutions*. Other l_p norms give a *Chebyshev solution* ($p = \infty$) and a l_1 solution. In the general case the objective function $\|Ax - b\|$ is typically nondifferentiable at the optimal solutions. Therefore a characterization and computing of $\|\cdot\|$-approximate soutions is hard (see for example Watson [25]).

Let $A \in \mathscr{R}_n^{m \times n}$. The *basic solutions* of $Ax = b$ are solutions of square subsystems of the original system, corresponding to nonsingular submatrices $A_J \times \mathscr{R}_n^{n \times n}$ of A, where $J = \{i_1, ..., i_n\}$ is a subset of $\{1, 2, ..., m\}$. The rows of the submatrix A_J correspond to rows of A with indices in the set J. Analogously we define b_J. Let

$$\mathscr{J}(A) = \{J \colon \det(A_J) \neq 0\}. \tag{3}$$

Then

$$\mathscr{C} \equiv \mathscr{C}(A, b) = \operatorname{conv}\{A_J^{-1} b_J \colon J \in \mathscr{J}(A)\} \tag{4}$$

is the *convex hull of the basic solutions* of $Ax = b$. Recently, Ben-Tal and Teboulle [3] have shown that, for $A \in \mathscr{R}_n^{m \times n}$ and $1 \leqslant p < \infty$, all l_p-approximate solutions of (1) lie in the convex hull $\mathscr{C}$. For $p = \infty$ there is an l_∞-approximate solution in $\mathscr{C}$. We recall that Chebyshev and l_1-approximate solutions are not unique in the general case.

The result of Ben-Tal and Teboulle [3] was stated for isotone functions, of which l_p norms can be considered a special case, and $A \in \mathscr{R}_n^{m \times n}$. Namely, they have dealt with the problem

$$\min_{x \in \mathscr{R}^n} f(|Ax - b|), \tag{5}$$

where $|x|$ denotes a vector whose elements are moduli of elements of x and f is an isotone function. A continuous function $f \colon \mathscr{R}_+^m \to \mathscr{R}$ is *isotone* if (compare Ortega and Rheinbold [17, Definition 2.4.3, p. 52])

$$0 \leqslant x \leqslant y \qquad \text{implies} \quad f(x) \leqslant f(y),$$

where the inequalities between vectors are interpreted componentwise. Moreover, if for

$$0 \leqslant x \leqslant y, \qquad f(x) = f(y) \qquad \text{implies} \quad x = y,$$

then f is *strictly isotone*. The l_p norm is strictly isotone for $1 < p < \infty$, because

$$0 \leqslant x \leqslant y \qquad \text{and} \qquad \|x\|_p = \|y\|_p \qquad \text{implies} \quad x = y.$$

Ben-Tal and Teboulle [3] have proven that if f is isotone then for an arbirary $A \in \mathscr{R}_n^{m \times n}$ there exists one solution of the optimization problem (5) which lies in the convex hull $\mathscr{C}$ of the basic solutions of $Ax = b$. Moreover, if f is strictly isotone then every solution lies in $\mathscr{C}$. This approach was inspired by the paper of Berg [4] who has proven this for the l_2 norm. In this paper we extend the result of Ben-Tal and Teboulle to slightly more general functions than f in (5). This leads to the corollary that every solution of (1) lies in $\mathscr{C}$ for norms which we call strictly orthant–monotonic norms.

The above results are a generalization of the result of Levitan and Lynn [12] who have considered the l_p norms and A satisfying

$$A \in \mathscr{R}^{(n+1) \times n}, \qquad \text{rank } A = n. \tag{6}$$

Levitan and Lynn have expressed explicitly the l_p-approximate solutions of $Ax = b$ as convex combinations of the basic solutions for the case (6). These formulae have been presented once more by Miao and Ben-Israel [15]. The problem (1) for A satisfying (6) or

$$A \in \mathscr{R}^{m \times n}, \qquad \text{rank } A = m - 1 \tag{7}$$

is equivalent to solving some consistent linear system (see Section 3). Using this system we explicitly express the solutions of (1) for some norms as convex combinations of the basic solutions. Additionally, we explain why this expression is not valid for any arbitrary norm. We will show which properties of a norm decide that a solution of (1) lies in $\mathscr{C}$. For this purpose we investigate the properties of the orthant–monotonic norms (see Section 2).

We also deal with the weighted approximate solutions of $Ax = b$. Namely, we consider

$$\min_x \|D(Ax - b)\|, \tag{8}$$

where D is a nonsingular diagonal matrix. The problem (8) is connected with weighted projections (see Hanke and Neumann [10], O'Leary [11], and Stewart [23]; compare with Forsgren [7] and Miao and Ben-Israel [16]). We give a generalization of some result of Miao and Ben-Israel [16] concerning the problem (8) with l_p norms. Our generalization holds for some strictly orthant–monotonic norms.

2. ORTHANT–MONOTONIC NORMS

Let $\|\cdot\|$ be a norm in $\mathscr{R}^m$. We recall that we have assumed only that a norm is *weakly homogenous* (see (2)). The norm $\|\cdot\|^*$ dual to $\|\cdot\|$ is defined by

$$\|v\|^* = \max_{\|y\| \leqslant 1} y^T v. \tag{9}$$

Let $v \neq 0$ and let v^* be a vector for which the maximum (9) is reached

$$v^T v^* = \|v\|^*, \qquad \|v^*\| = 1.$$

The vector v^* is called $\|\cdot\|$-*dual vector* to v (see Sreedharan [20]). This definition is slightly different from the definition of a dual vector pair in Bauer *et al.* [1]. Namely, they call (x, y) a *dual vector pair* with respect to a norm $\|\cdot\|$ if

$$y^T x = \|x\| \; \|y\|^*. \tag{10}$$

Thus if $x \neq 0$ then $x/\|x\|$ is a $\|\cdot\|$-dual vector to y. Therefore for x, $\|x\| = 1$, a vector pair (x, y) is dual if and only if x is a $\|\cdot\|$-dual vector to y, $x = y^*$.

In the general case a $\|\cdot\|$-dual vector is not unique. We now recall the formulae for the dual vectors with respect to the l_p norm, $1 \leqslant p \leqslant \infty$. For $1 < p < \infty$ the l_p-dual vector $v^* = [v_1^*, ..., v_m^*]^T$ to a nonzero $v = [v_1, ..., v_m]^T \in \mathscr{R}^m$ is unique and it has the elements

$$v_j^* = \operatorname{sgn}(v_j)(|v_j|/\|v\|_q)^{q-1}, \tag{11}$$

where $1/p + 1/q = 1$. A l_1-dual vector has the elements

$$v_j^* = \begin{cases} \operatorname{sgn}(v_j)\, g_j & \text{if} \quad |v_j| = \|v\|_\infty, \\ 0 & \text{if} \quad |v_j| < \|v\|_\infty, \end{cases} \tag{12}$$

where $g_j \geqslant 0$ and $\sum_j g_j = 1$. For $p = \infty$ we have

$$v_j^* = \begin{cases} \operatorname{sgn}(v_j) & \text{if} \quad v_j \neq 0, \\ h_j & \text{if} \quad v_j = 0, \end{cases} \tag{13}$$

where $|h_j| \leqslant 1$.

In connection with some properties of norms which are needed in the next section we introduce auxiliary definitions. Let

$$\mathscr{J}(v) \overset{\text{df}}{=} \{j \colon v_j \neq 0\}, \qquad v \in \mathscr{R}^m.$$

We say that a norm $\|\cdot\|$ in $\mathscr{R}^m$ has *property P*1 if for every nonzero $v = [v_1, ..., v_m]^T \in \mathscr{R}^m$ and for every $\|\cdot\|$-dual vector $v^* = [v_1^*, ..., v_m^*]^T$ to v we have

$$v_j v_j^* \geqslant 0, \qquad j = 1, ..., m. \tag{14}$$

If the inequalities (14) and

$$v_j^* = 0 \qquad \text{for} \quad j \notin \mathscr{J}(v) \tag{15}$$

hold for every v^* then we say that the norm $\|\cdot\|$ has *property P*2. If a norm has the property P1 and we have

$$v_j^* = 0 \qquad \text{if and only if} \quad v_j = 0 \tag{16}$$

then the norm has *property P*3. Immediately from (11), (12), and (13) it follows that the l_p norm has the property P1 for $1 \leqslant p \leqslant \infty$. Moreover, the l_p norm has property P2 for $1 \leqslant p < \infty$ and the property P3 for $1 < p < \infty$. We will characterize norms which have the property P1, P2, or P3.

The l_p norms are orthant–monotonic ones. A norm $\|\cdot\|$ in $\mathscr{R}^m$ is *orthant–monotonic* if for all vectors $x = [x_1, ..., x_m]^T$, $y = [y_1, ..., y_m]^T$ (see Gries [8], Gries and Stoer [9], Loizou [13])

$$|x| \leqslant |y| \qquad \text{and} \qquad y_j x_j \geqslant 0 \quad (j = 1, ..., m) \text{ implies } \|x\| \leqslant \|y\|.$$

Increasingly important absolute norms are orthant–monotonic, but not every orthant–monotonic norm is absolute. A norm $\|\cdot\|$ is absolute if

$$\|x\| = \|\,|x|\,\|.$$

Gries [8] has given an example of an orthant–monotonic norm in $\mathscr{R}^2$, which is not absolute,

$$\|x\| = \begin{cases} \|x\|_\infty & \text{if} \quad x_1 \geqslant 0, \\ \|x\|_2 & \text{if} \quad x_1 \leqslant 0. \end{cases}$$

It is known that a norm is absolute if and only if it is monotonic, i.e. (see Bauer, *et al.* [1]),

$$|x| \leqslant |y| \qquad \text{implies} \quad \|x\| \leqslant \|y\|.$$

The orthant–monotonic norms have interesting properties. We now recall a theorem which characterizes these norms (see Gries [8]).

THEOREM 1 (Gries). *Let* $\|\cdot\|$ *be a norm in* $\mathscr{R}^m$. *Then the following statements are equivalent*

(i) $\|\cdot\|$ *is orthant–monotonic,*

(ii) $\|\cdot\|^*$ *is orthant–monotonic,*

(iii) $v_j v_j^* \geqslant 0$ $(j=1, ..., m)$ *for every* $\|\cdot\|$*-dual vector* v^* *to an arbitrary* v, *i.e., the norm* $\|\cdot\|$ *has the property* $P1$.

Remark. The last statement of Theorem 1 is formulated in Gries [8] for a dual vector pair (see (10)). Namely, a norm $\|\cdot\|$ is orthant–monotonic if and only if

$$y^T x = \|x\| \ \|y\|^* \qquad \text{implies} \quad y_j x_j \geqslant 0 \qquad (j=1, ..., m). \tag{17}$$

Other characterizations of orthant–monotonic norms are given in de Sá and Sodupe [19]. We recall one of them.

THEOREM 2 (de Sá and Sodupe). *A norm* $\|\cdot\|$ *in* $\mathscr{R}^m$ *is orthant–monotonic if and only if for any* $u \neq 0$ *there exists* $v \neq 0$ *such that* $v^T u = \|u\| \ \|v\|^*$ *and* $v_i = 0$ *if* $u_i = 0$.

Theorem 2 implies that a norm $\|\cdot\|$ is orthant–monotonic if and only if for any $u \neq 0$ there exists $v \neq 0$ such that $v/\|v\|^*$ is $\|\cdot\|^*$-dual to u and $v_i = 0$ if $u_i = 0$. Therefore a norm $\|\cdot\|$ is orthant–monotonic if and only if for any $v \neq 0$ there exists $\|\cdot\|$-dual vector v^* such that $v_i^* = 0$ if $v_i = 0$ (see Theorems 1 and 2).

Let us define a strictly orthant–monotonic norm. We say that a norm $\|\cdot\|$ in $\mathscr{R}^m$ is *strictly orthant–monotonic* if it is orthant–monotonic and

$$|x| \leqslant |y|, \qquad x_j \, y_j \geqslant 0 \quad (j=1, ..., m), \qquad \|x\| = \|y\| \qquad \text{implies} \quad x = y.$$

This definition is analogous to the definition of a strictly isotone function. We now prove a characterization of a strictly orthant–monotonic norm.

THEOREM 3. *A norm* $\|\cdot\|$ *in* $\mathscr{R}^m$ *is strictly orthant–monotonic if and only if it has the property* $P2$.

Proof. Let $\|\cdot\|$ be strictly orthant–monotonic. Let v^* be an arbitrary $\|\cdot\|$-dual to a nonzero v. We define $\tilde{v}^* = [\tilde{v}_1^*, ..., \tilde{v}_m^*]^T$,

$$\tilde{v}_j^* = \begin{cases} 0 & \text{if} \quad j \notin \mathscr{J}(v), \\ v_j^* & \text{if} \quad j \in \mathscr{J}(v). \end{cases}$$

Then

$$|\tilde{v}^*| \leqslant |v^*|, \qquad \tilde{v}_j^* v_j^* \geqslant 0, \qquad \|\tilde{v}^*\| \leqslant \|v^*\| = 1,$$

$$\|v\|^* = \sum_j v_j v_j^* = \sum_j v_j \tilde{v}_j^* \leqslant \|v\|^* \, \|\tilde{v}^*\| \leqslant \|v\|^* \, \|v^*\| = \|v\|^*.$$

Therefore $\|\tilde{v}^*\| = \|v^*\| = 1$. This implies $\tilde{v}^* = v^*$. Therefore $v_j^* = 0$ for $j \notin \mathscr{J}(v)$. Thus the condition (15) is fulfilled for every dual vector v^*. The condition (14) also is satisfied because the norm is orthant–monotonic.

Let now a norm $\|\cdot\|$ have the property P2. Then the norm $\|\cdot\|$ is orthant–monotonic because of the condition (14) (see Theorem 1). Let x be a nonzero vector and let $y \in \mathscr{R}^m$ satisfy

$$0 \leqslant |x| \leqslant |y|, \qquad x_j y_j \geqslant 0 \quad \text{for every } j, \qquad \|x\| = \|y\|. \tag{18}$$

Let $z = [z_1, ..., z_m]^T$ be a $\|\cdot\|^*$-dual vector to x. Then we have

$$x^T z = \|x\|, \qquad \|z\|^* = 1, \qquad z_j x_j \geqslant 0 \quad \text{for each } j \tag{19}$$

because the dual norm $\|\cdot\|^*$ is also orthant–monotonic (see Theorem 1). On the other hand, the relations (19) mean that $x/\|x\|$ is a $\|\cdot\|$-dual vector to z. Therefore

$$x_j = 0 \qquad \text{for} \quad j \notin \mathscr{J}(z)$$

since we have the property (15). Thus (see (18) and (19))

$$\|x\| = \|x\| \, \|z\|^* = x^T z \leqslant y^T z \leqslant \|y\| \, \|z\|^* = \|y\|.$$

This means that $y/\|y\|$ is some $\|\cdot\|$-dual vector to z. Therefore $y_j = 0$ for $j \notin \mathscr{J}(z)$. Consequently we have

$$\sum_{j \in \mathscr{J}(z)} x_j z_j = \sum_{j \in \mathscr{J}(z)} y_j z_j = \sum_{j \in \mathscr{J}(z)} |x_j| \, |z_j| = \sum_{j \in \mathscr{J}(z)} |y_j| \, |z_j|,$$

$$\sum_{j \in \mathscr{J}(z)} |z_j|(|y_j| - |x_j|) = 0.$$

Hence $x_j = y_j$ for $j \in \mathscr{J}(z)$ because $z_j \neq 0$ for $j \in \mathscr{J}(z)$, $x_j y_j \geqslant 0$, and $|y_j| \geqslant |x_j|$. This completes the proof because $x_j = y_j = 0$ for $j \notin \mathscr{J}(z)$, so $x = y$. ∎

We now give a sufficient condition for norms which have the property P3. For this reason we recall the definition of the *smoothness of a norm.* A norm $\|\cdot\|$ is smooth if and only if at every point of the unit norm there is exactly one hyperplane supporting the closed unit ball $\mathscr{B} = \{x \in \mathscr{R}^n: \|x\| \leqslant 1\}$. Now $\mathscr{H}$ is a hyperplane supporting $\mathscr{B}$ at x_1, $\|x_1\| = 1$, if and only if there exists $y \in \mathscr{R}^n$ such that $\mathscr{H} = \{z \in \mathscr{R}^n: z^T y = 1\}$ with $y^T x_1 = 1$ and $\|y\|^* = 1$. In other words, a norm $\|\cdot\|$ is smooth if and only if there exists exactly one $\|\cdot\|^*$-dual vector to x_1 (see for example Sreedharan [21]).

THEOREM 4. *Let a norm $\|\cdot\|$ be smooth strictly orthant–monotonic. Then $\|\cdot\|$ has the property $P3$.*

Proof. We recall that a norm $\|\cdot\|$ is smooth if and only if $\|\cdot\|^*$ is strictly convex. Therefore a $\|\cdot\|^*$-dual vector to a nonzero vector is unique (compare Sreedharan [22, Lemma 2.1]).

Let v^* be an arbitrary $\|\cdot\|$-dual vector to nonzero $v \in \mathscr{R}^m$. Then

$$v_j = 0 \qquad \text{implies} \quad v_j^* = 0 \tag{20}$$

because the norm is strictly orthant–monotonic (see Theorem 3)). The dual norm is orthant–monotonic and strictly convex. Therefore a $\|\cdot\|^*$-dual vector v' to v^* is unique and v' has the property (15), i.e., $v_j^* = 0$ implies $v'_j = 0$ (see Theorem 2). However, it is easy to verify that by the uniqueness we have $v' = v/\|v\|^*$. Therefore $v_j^* = 0$ implies $v_j = 0$. This completes the proof because of (20). ∎

Gries and Stoer [9] have proven the following very useful property of orthant–monotonic norms.

THEOREM 5 (Gries, Stoer). *Let $\|\cdot\|$ be an orthant–monotonic norm in $\mathscr{R}^m$ and let $x = [x_1, ..., x_m]^T \geqslant 0$, $y = [y_1, ..., y_m]^T \geqslant 0$ be nonzero vectors with*

$$x_j = 0 \qquad \text{if and only if} \quad y_j = 0. \tag{21}$$

Then there exists a diagonal matrix D with positive diagonal elements such that

$$(Dx)^T (D^{-1}y) = x^T y = \|D^{-1}y\| \, \|Dx\|^*. \tag{22}$$

Theorem 5 was proved by Stoer and Witzgall [24] for $x > 0$ and $y > 0$.

If x and y satisfy (21) and $x_j y_j \geqslant 0$ for every j then $\tilde{x} = \tilde{D}x$ and $\tilde{y} = \tilde{D}y$ are nonnegative for

$$\tilde{D} = \operatorname{diag}(\tilde{d}_j),$$

where $\tilde{d}_j = \operatorname{sgn}(x_j)$ for $x_j \neq 0$ and $\tilde{d}_j = 1$ otherwise. Moreover, $\tilde{x}^T \tilde{y} = x^T y$. Therefore from Theorem 5 we obtain immediately the following corollary.

COROLLARY 6. *Let a norm* $\|\cdot\|$ *be orthant–monotonic, let nonzero vectors* $x, y \in \mathscr{R}^m$ *satisfy* (21), *and let* $x_j y_j \geqslant 0$ *for every* j. *Then there exists a nonsingular diagonal matrix* $D = \operatorname{diag}(d_j)$ *such that* $d_j x_j \geqslant 0$ *and* $D^{-1}y/\|D^{-1}y\|$ *is a* $\|\cdot\|$*-dual vector to* Dx.

We will use Corollary 6 in the next section to investigate the problem (8).

3. OVERDETERMINED LINEAR SYSTEMS

It is well known that a vector $x \in \mathscr{R}^n$ is a solution of (1) if and only if there exists $v \in \mathscr{R}^m$ such that (see for example Watson [25, Theorem 1.7])

$$\|Ax - b\| = -v^T b, \qquad v^T A = 0, \quad \|v\|^* = 1. \tag{23}$$

Let x be a solution of (1). From (23) we obtain $\|Ax - b\| \equiv \delta = v^T(Ax - b) = -v^T b$. Thus the vector $(Ax - b)/\delta$ is some $\|\cdot\|$-dual vector to v. Therefore there exists a $\|\cdot\|$-dual vector v^* to v such that x is a solution of the consistent linear system

$$Ax = b - (v^T b) v^*. \tag{24}$$

Unfortunately, in the general case not every $\|\cdot\|$-dual vector v^* to v gives by (24) a solution of (1) because $b + \delta v^*$ has to belong to the linear subspace spanned by columns of A.

We now prove a theorem which shows that the problem (1) for A satisfying (7) is equivalent to solving some consistent linear system. We note that in this case $\dim \ker A^T = 1$.

THEOREM 7. *Let* $A \in \mathscr{R}^{m \times n}_{m-1}$ *and let* $\|\cdot\|$ *be an arbitrary norm in* $\mathscr{R}^m$. *Then* x *solves* (1) *if and only if* x *is a solution of the consistent linear system*

$$Ax = b - (w^T b) w^*, \tag{25}$$

where the vector w *satisfies*

$$w^T A = 0, \qquad \|w\|^* = 1, \qquad w^T b < 0, \tag{26}$$

and w^* *is a* $\|\cdot\|$*-dual vector to* w. *Moreover,* $\delta = -w^T b$.

Proof. Let w satisfy (26). Then $b - (w^T b) w^* \in \operatorname{span}\{w\}^{\perp}$ because $w^T[b - (w^T b) w^*] = 0$. This implies that $b - (w^T b) w^*$ belongs to the linear subspace spanned by columns of A because $\ker A^T = \operatorname{span}\{w\}$. Therefore the linear system (25) is consistent for every $\|\cdot\|$-dual vector w^*.

Let x be a solution of (1). Then there exists v satisfying (23). Thus the vector $(Ax-b)/\delta$ is some $\|\cdot\|$-dual vector v^* to v and consequently (24) holds. The vector $v \in \text{span}\{w\}$ because $\dim \ker A^T = 1$. Thus $v = \alpha w$ for some $\alpha > 0$ since $w^T b < 0$ and $v^T b < 0$. Therefore $v = w$ because $\|v\|^* = \|w\|^*$. This implies that x is a solution of (25). Moreover, $\delta = -w^T b$.

Let now x be a solution of (25). Then $\|Ax - b\| = -w^T b = \delta$ which means that x is a solution of (1). This completes the proof. ∎

Theorem 7 was proved by Ziętak [26] for A satisfying (6) and strictly homogeneous norms (compare Cheney [5, p. 41], Levitan and Lynn [12], Meicler [14], Miao and Ben-Israel [15], and Sreedharan [20]). We stress that the linear system (25) is consistent for every w^*, i.e., every w^* determines a solution of (1). Moreover, for each solution x of (1) we can choose such a w^* that x is a solution of the consistent system (25).

Let $A \in \mathscr{R}_n^{(n+1)\times n}$ and let $u = [u_1, ..., u_{n+1}]^T$ have the elements

$$u_j = (-1)^j \det A_j, \tag{27}$$

where A_j is obtained from A by deletion of the jth row. This notation is slightly different from the definition of A_J in (3). It is easily seen that

$$\det[A, b] = (-1)^{n+1} u^T b. \tag{28}$$

The matrix $[A, b]$ is nonsingular and $u^T b \neq 0$ because we have assumed that b is not in the subspace spanned by columns of A and rank $A = n$. Moreover, we have $u^T A = 0$. It is well known that the least squares solution $x^{(2)}$ of $Ax = b$ is the unique solution of the so-called *normal equation* $A^T A x = A^T b$. This implies that

$$\ker A^T = \text{span}\{u\} = \text{span}\{r^{(2)}\},$$

where $r^{(2)} = Ax^{(2)} - b$. Let

$$w = \beta u / \|\beta u\|^*, \qquad \text{where} \quad \beta = -\text{sgn}(u^T b). \tag{29}$$

Then w satisfies the conditions (26).

The submatrices A_j are nonsingular for each $j \in \mathscr{J}(u)$. The basic solutions for the system $Ax = b$, $A \in \mathscr{R}_n^{(n+1)\times n}$, are

$$\{A_j^{-1} b^{(j)} : j \in \mathscr{J}(u)\}, \tag{30}$$

where $b^{(j)}$ denotes a vector obtained from b by deletion of its jth element. Levitan and Lynn [12] and Miao and Ben-Israel [15] have expressed the l_p-approximate solutions of (1) for $1 < p < \infty$ as convex combinations of the basic solutions (30). Moreover, they have shown that some solutions of (1) for $p=1$ and $p=\infty$ are also convex combinations of the basic solutions. Namely, for $p=1$ the solution

$$x = A_l^{-1} b^{(l)}, \qquad \text{where the index } l \text{ is such that } |u_l| = \max_j |u_j|, \tag{31}$$

lies in $\mathscr{C}$. On the other hand, for $p=\infty$ the solution

$$x = \sum_{j \in \mathscr{J}(u)} \frac{|u_j|}{\|u\|_1} A_j^{-1} b^{(j)} \tag{32}$$

is in $\mathscr{C}$. Let $\mu_j^{(\infty)} = w_j w_j^*$ for every j, where w is determined as in (29) and w^* is a l_∞-dual vector to w. Then the vector (32) has the form (see (13))

$$x = \sum_{j \in \mathscr{J}(u)} \mu_j^{(\infty)} A_j^{-1} b^{(j)}.$$

Analogously, we verify that the vector (31) has, in fact, the form

$$\sum_{j \in \mathscr{J}(u)} \mu_j^{(1)} A_j^{-1} b^{(j)},$$

where $\mu_j^{(1)} = w_j w_j^*$ with a l_1-dual vector w^* such that in (12) $g_l = 1$ and $g_j = 0$ for other j. The formulae (31) and (32) suggest the following generalization.

THEOREM 8. *Let* $\|\cdot\|$ *be an arbitrary norm,* $A \in \mathscr{R}_n^{(n+1)\times n}$, *and let* $w^* = [w_1^*, ..., w_{n+1}^*]^T$ *be an arbitrary* $\|\cdot\|$*-dual vector to* $w = [w_1, ..., w_{n+1}]^T$ *determined as in* (29). *Then*

$$x = \sum_{j \in \mathscr{J}(w)} w_j w_j^* A_j^{-1} b^{(j)} \tag{33}$$

is a solution of (1) *if and only if*

$$w_j^* = 0 \qquad \textit{for} \quad j \notin \mathscr{J}(w). \tag{34}$$

Remark. We assume that a norm $\|\cdot\|$ is such that there exists a $\|\cdot\|$-dual vector w^* to w satisfying (34).

Proof. We denote $y^{(j)} \equiv [y_1^{(j)}, ..., y_{n+1}^{(j)}]^T = AA_j^{-1}b^{(j)} - b$ for $j \in \mathscr{J}(w)$. Then for $j \in \mathscr{J}(w)$ we have (compare the proof of Lemma 3.1 in Miao and Ben-Israel [15])

$$(-1)^{n+1-j} \det[A, b] = (b_j - \alpha_j A_j^{-1} b^{(j)}) \det A_j = -y_j^{(j)} \det A_j,$$

where α_j denotes the jth row of A, and b_j is the jth element of b. Therefore for $j \in \mathcal{J}(w)$ (see (28))

$$y_j^{(j)} = -u^T b/u_j = -w^T b/w_j$$

and consequently

$$y_i^{(j)} = \begin{cases} 0 & \text{if} \quad i \neq j, \\ -w^T b/w_j & \text{if} \quad i = j. \end{cases}$$

Moreover, we have

$$\sum_{j \in \mathcal{J}(w)} w_j w_j^* = 1 \tag{35}$$

because $\|w\|^* = 1$. Let x be determined as in (33). Then

$$z \equiv Ax - b = \sum_{j \in \mathcal{J}(w)} w_j w_j^* (AA_j^{-1} b^{(j)} - b) = \sum_{j \in \mathcal{J}(w)} w_j w_j^* y^{(j)}. \tag{36}$$

Therefore ($z = [z_1, \ldots, z_{n+1}]^T$)

$$z_j = \begin{cases} 0 & \text{if} \quad j \notin \mathcal{J}(w), \\ -(w^T b) w_j^* & \text{if} \quad j \in \mathcal{J}(w). \end{cases}$$

Thus $z = -(w^T b) w^*$ if and only if w^* satisfies (34). This completes the proof because x solves (1) if and only if $Ax - b = -(w^T b) w^*$ for some w^* (see Theorem 7). ∎

If

$$w_j w_j^* \geqslant 0 \qquad \text{for each } j \tag{37}$$

and (34) is fulfilled then the solution (33) is a convex combination of the basic solutions (30) because we have (35). Unfortunately, this does not hold in the general case for an arbitrary norm. The condition (37) is satisfied for orthant–monotonic norms (see Theorem 1).

The coefficients $w_j w_j^*$ in (33) depend on b in the general case because of the assumption $w^T b < 0$. However, they do not depend on b in (31) and (32). The l_1 and l_∞ norms are strictly homogenous. It is easily seen that if a norm is strictly homogenous then we can omit the assumption $w^T b < 0$ in (26). It is connected with the following property of dual vectors. Let $\|\cdot\|$ be a strictly homogeneous norm. Then a vector z is $\|\cdot\|$-dual to $y = -w$ if and only if $z = -w^*$, where w^* is a $\|\cdot\|$-dual vector to w. Thus $(w^T b) w^* = (y^T b) z$. Therefore for strictly homogeneous norms we can choose common

w for all b in Theorem 7 and we have $\delta = |w^T b|$. Unfortunately, we cannot omit the assumption $w^T b < 0$ in the general case for weakly homogenous norms.

We now give another formula for solutions of (1). This formula is valid for arbitrary norms and A satisfying (6).

THEOREM 9. *Let the assumptions of Theorem* 8 *be satisfied. Then x is a solution of* (1) *if and only if it has the form*

$$x = \frac{\beta}{\|\beta u\|^*} \sum_{j=1}^{n+1} (-1)^j w_j^* (\operatorname{adj} A_j)\, b^{(j)}, \tag{38}$$

where u and β are determined as in (27) *and* (29), *respectively*, *and* adj A_j *denotes the adjoint of A_j.*

Proof. The consistent linear system (25) can be written in the form

$$[A, b] \begin{bmatrix} x \\ -1 \end{bmatrix} = -(w^T b) w^*.$$

A vector x is a solution of (1) if and only if x solves this system for the appropriate dual vector w^*. Solving this system by Cramer's rule we obtain immediately (38). This completes the proof. ∎

If adj A_j is a nonsingular matrix then

$$\frac{\beta}{\|\beta u\|^*} \operatorname{adj} A_j = (-1)^j w_j A_j^{-1}$$

and consequently (see (38))

$$(-1)^j \frac{\beta w_j^*}{\|\beta u\|^*} \operatorname{adj} A_j = w_j w_j^* A_j^{-1}.$$

This explains why the property (34) is crucial if we want to express the solution of (1) as a combination of the basic solutions. Moreover, we have the following corollary.

COROLLARY 10. *Let the assumptions of Theorem* 8 *be satisfied. Then the vector* (38) *has the form* (33) *if and only if* (34) *holds.*

From Theorems 2, 3, and 8 we obtain immediately the following corollary which is one of the main results of this section.

Corollary 11. *Let* $A \in \mathscr{R}_n^{(n+1)\times n}$ *and let* w *be determined as in* (29).

(i) *There exists a solution of* (1) *which lies in the convex hull* $\mathscr{C}$ *of the basic solutions* (30) *if and only if a norm is orthant–monotonic. Moreover, this solution has the form* (33) *for appropriate* w^*.

(ii) *Every solution of* (1) *is a convex combination of the basic solutions* (30) *if and only if a norm is strictly orthant–monotonic. Moreover, in this case all solutions have the form* (33).

A matrix $A \in \mathscr{R}_n^{(n+1)\times n}$ is a Haar matrix if every $\det A_j$ is different from zero. For a Haar matrix the l_∞-dual vector w^* to w is unique. In this case the unique Chebyshev solution of (1) has the form (32) and it lies in $\mathscr{C}$. If $A \in \mathscr{R}_n^{(n+1)\times n}$ is not a Haar matrix then the solution of (1) is not unique and the only l_∞-dual vector w^* satisfying (34) has the elements

$$w_j^* = \begin{cases} \operatorname{sgn}(w_j) & \text{if } \ j \in \mathscr{J}(w), \\ 0 & \text{if } \ j \notin \mathscr{J}(w). \end{cases}$$

For this l_∞-dual vector w^* the vector (33) has the form (32). Therefore the vector (32) is the only Chebyshev solution of (1) which lies in $\mathscr{C}$. On the other hand, it is well known that this vector is the *strict Chebyshev solution* and it is the limit of the l_p-approximate solutions when p tends to ∞ (for the definition and properties of the strict Chebyshev solution of an over-determined linear system see for example Descloux [6] and Rice [18]).

As we have mentioned in the Introduction, Levitan, Lynn, Miao, and Ben-Israel have expressed the l_p-approximate solutions of $Ax = b$ for $A \in \mathscr{R}_n^{(n+1)\times n}$ as the convex combination of the basic solutions. Corollary 11 generalizes their result and answers the question when a solution lies in $\mathscr{C}$. The same approach can be used to a generalization of some results which are known for minimum l_2-norm l_p-approximate solutions for $A \in \mathscr{R}_{m-1}^{m\times n}$ (see Section 4 in Miao and Ben-Israel [15]). This possible generalization will be considered in another paper.

If an isotone function f is such that $f(|x|)$ is a norm in $\mathscr{R}^m$ then this norm is absolute. Therefore the problem (5), for an isotone function f being a norm, is the problem (1) for an absolute norm. Therefore the result of Ben-Tal and Teboulle, concerning the problem (5), is valid also for the problem (1) with absolute norms. On the other hand, we have Corollary 11 and we know that an absolute norm is orthant–monotonic. This suggests that the result of Ben-Tal and Teboulle holds for slightly more general functions f than isotone ones. Indeed, we show that the proof of their Theorem 2.2 in [3] is true also for the problem

$$\min_{x \in \mathscr{R}^n} g(Ax - b), \tag{39}$$

where the continuous function $g\colon \mathscr{R}^m \to \mathscr{R}$ satisfies

$$|x| \leqslant |y|, \qquad x_j y_j \geqslant 0 \quad (j=1, ..., m) \text{ implies } g(x) \leqslant g(y). \tag{40}$$

We say that g, satisfying (40), is *orthant–isotone*. If additionally we have that

$$g(x) = g(y) \qquad \text{implies} \quad x = y, \tag{41}$$

then g is *strictly orthant–isotone*. From (40) we obtain $g(0) \leqslant g(y)$ for every $y \in \mathscr{R}^m$. We now prove a generalization of Theorem 2.2 from [3].

THEOREM 12. *Let g be orthant–isotone and let $A \in \mathscr{R}_n^{m \times n}$. Then there exists a solution of* (39) *which lies in the convex hull $\mathscr{C}$ of the basic solutions of $Ax = b$. Moreover, if g is strictly orthant–isotone, then every solution of* (39) *lies in $\mathscr{C}$.*

Proof. The proof of the theorem is exactly the same as the proof of Theorem 2.2 in Ben-Tal and Teboulle [3] because a stronger relation than (7) in [3] is true in fact. Therefore we give only the second part of the proof.

Let $x \in \mathscr{R}^n$ and let $x = s + d$, where the statement of s and d is given in the proof of Theorem 2.2 in [3]. Then $s \in \mathscr{C}$ and we have (see (7) in [3])

$$|Ax - b| = |As - b| + |Ad|. \tag{42}$$

However, a stronger relation than (42) holds because the ith elements ($i = 1, ..., m$) of the vectors $Ax - b$ and $As - b$ have a common sign. More precisely, we have

$$[(Ax-b)_i][(As-b)_i] > 0 \qquad \text{or} \qquad (Ax-b)_i = (As-b)_i = 0 \tag{43}$$

where $(Ax - b)_i$ denotes the ith element of the vector $Ax - b$. This follows immediately from the definitions of s and d (see Ben-Tal and Teboulle [3]). Therefore $g(Ax - b) \geqslant g(As - b)$ because g is orthant–isotone. Consequently we obtain

$$\min_{x \in \mathscr{R}^n} g(Ax - b) = \min_{s \in \mathscr{C}} g(As - b).$$

Thus the first part of the theorem is proven.

Let now g be strictly orthant–isotone and let x be any solution of (39). Then $x = s + d$ as above. From the first part of the theorem and the fact that x is a solution of (39) it follows that $g(As - b) = g(Ax - b)$. Since g is strictly orthant–isotone we must have $As - b = Ax - b$ and $Ad = 0$ by (42) and (43). Because rank of A is n we obtain $d = 0$ which implies $x = s \in \mathscr{C}$. This completes the proof. ∎

From Theorem 12 we obtain the following corollary.

COROLLARY 13. *Let $A \in \mathscr{R}_n^{m \times n}$. If a norm $\|\cdot\|$ is orthant–monotonic then there is a solution of* (1) *in the convex hull $\mathscr{C}$ of the basic solutions* (4). *Moreover, if $\|\cdot\|$ is strictly orthant–monotonic then every solution lies in $\mathscr{C}$.*

Let us return to the problem (8) with an arbitrary matrix $A \in \mathscr{R}_n^{m \times n}$. Miao and Ben-Israel [16] have considered relations between the sets

$$\mathscr{H}_+ = \{(A^T D A)^{-1} A^T D b \colon D \in \mathscr{D}_+\}, \tag{44}$$

$$\mathscr{H}_+^{(\|\cdot\|)} = \bigcup_{D \in \mathscr{D}_+} \{x \colon \arg\min_x \|D(Ax - b)\|\}, \tag{45}$$

where $\mathscr{D}_+$ denotes the set of all nonsingular diagonal matrices with positive diagonal elements. The set $\mathscr{H}_+$ is the set of all solutions of the weighted least squares problems for all $D \in \mathscr{D}_+$

$$\bigcup_{D \in \mathscr{D}_+} \{x \colon \arg\min_x \|D^{1/2}(Ax - b)\|_2\}.$$

Therefore $\mathscr{H}_+^{(l_2)} = \mathscr{H}_+$. Miao and Ben-Israel [16] have shown that for $A \in \mathscr{R}_n^{m \times n}$ we have the following relations

$$\mathscr{H}_+^{(l_2)} = \mathscr{H}_+^{(l_p)}, \qquad 1 < p < \infty, \tag{46}$$

$$\mathscr{H}_+^{(l_2)} \subseteq \mathscr{H}_+^{(l_1)}. \tag{47}$$

We now extend this.

Let now $A \in \mathscr{R}^{m \times n}$. Analogously as (45), we define $\mathscr{H}^{(\|\cdot\|)}$ for $\mathscr{D}$ which is the set of all nonsingular diagonal matrices D

$$\mathscr{H}^{(\|\cdot\|)} = \bigcup_{D \in \mathscr{D}} \{x \colon \arg\min_x \|D(Ax - b)\|\}. \tag{48}$$

We now prove that relations similar to (46) and (47) hold between $\mathscr{H}^{(\|\cdot\|)}$ for some class of orthant–monotonic norms, which contain the l_p norms as a subclass. We note that if rank $A < n$ then $\mathscr{H}^{(l_2)}$ defined as in (48) does not have the form (44).

THEOREM 14. *Let a norm $\|\cdot\|_{(1)}$ have the property $P3$, let a norm $\|\cdot\|_{(2)}$ have the property $P1$, and let A be an arbitrary matrix, $A \in \mathscr{R}^{m \times n}$. Then for every solution $\tilde{x}$ of the problem* (1) *with respect to the norm $\|\cdot\|_{(1)}$ there exists a nonsingular diagonal matrix $\tilde{D}$ such that $\tilde{x}$ is a solution of the problem*

$$\min_x \|\tilde{D}(Ax - b)\|_{(2)}. \tag{49}$$

Proof. Let $\tilde{x}$ be an arbitrary solution of (1) with respect to the norm $\|\cdot\|_{(1)}$. Then there exist a nonzero vector v and an $\|\cdot\|_{(1)}$-dual vector v^* to v such that $\tilde{x}$ is a solution of the consistent system (24). Moreover, (23) holds for $\tilde{x}$ and v. Since $\|\cdot\|_{(1)}$ has the property P3 we have (see (14) and (16))

$$v_j v_j^* > 0 \qquad \text{or} \qquad v_j = v_j^* = 0.$$

However, the norm $\|\cdot\|_{(2)}$ is orthant–monotonic. Therefore there exists a nonsingular diagonal matrix D such that (see Corollary 6)

$$\tilde{v}^* = D^{-1}v^*/\|D^{-1}v^*\|_{(2)}$$

is a $\|\cdot\|_{(2)}$-dual vector to $\tilde{v} = \|D^{-1}v^*\|_{(2)}\, Dv$ and $\|Dv\|^*_{(2)}\, \|D^{-1}v^*\|_{(2)} = 1$. We define

$$\tilde{b} = D^{-1}b, \qquad \tilde{A} = D^{-1}A.$$

Then we have

$$\|\tilde{A}\tilde{x} - \tilde{b}\|_{(2)} = -\tilde{v}^T\tilde{b}, \qquad \|\tilde{v}\|^*_{(2)} = 1, \qquad \tilde{v}^T\tilde{A} = 0.$$

This implies that $\tilde{x}$ is a solution of (49) for $\tilde{D} = D^{-1}$ because the conditions analogous to (23) are satisfied by $\tilde{v}$ for the problem (49). This completes the proof. ▮

From Theorem 14 we obtain immediately the following corollary.

COROLLARY 15. *Let the assumptions of Theorem* 14 *be satisfied. Then*

$$\mathscr{H}^{(\|\cdot\|)(1))} \subseteq \mathscr{H}^{(\|\cdot\|(2))}. \tag{50}$$

If additionally the norm $\|\cdot\|_{(2)}$ has the property $P3$ also then the equality holds in (50). *In particular, if a norm $\|\cdot\|$ is smooth strictly orthant–monotonic then*

$$\mathscr{H}^{(\|\cdot\|)} = \mathscr{H}^{(l_2)}.$$

Moreover, for any arbitrary orthant–monotonic norm $\|\cdot\|$ we have

$$\mathscr{H}^{(l_2)} \subseteq \mathscr{H}^{(\|\cdot\|)}.$$

We recall that the l_p norms, $1 < p < \infty$, have the property P3. Therefore it is not surprising that $\mathscr{H}_+^{(l_2)} = \mathscr{H}_+^{(l_p)}$ because in this case the problem (49) does not depend on the sign of the diagonal elements of D. We stress that the relations (46) and (47) were proven by Miao and Ben-Israel for matrices A of full column rank. However, in Theorem 14 the matrix A has arbitrary rank.

If $A \in \mathscr{R}_m^{m \times n}$ then the system $Ax = b$ is consistent, but its solution can be non-unique. Miao and Ben-Israel [16] have investigated the properties of the minimum l_p norm solutions of the consistent system $Ax = b$. Their results can be extended to the case of more general norms. A possible generalization will be presented in Ziętak [27].

Let G be a generalized inverse of A, $AGA = A$. A generalized inverse of A always exists and it is not unique unless A is nonsingular. A $\|\cdot\|$*-approximate generalized inverse* of A is such a generalized inverse G that (see for example Ziętak [26])

$$\|AGb - b\| = \min_{x \in \mathscr{R}^n} \|Ax - b\| \tag{51}$$

for all $b \in \mathscr{R}^m$. It is not guaranteed in the general case that such a matrix G exists. If $A \in \mathscr{R}_n^{(n+1) \times n}$ then a $\|\cdot\|$-approximate inverse exists and all approximate inverses are given explicitly, by very simple formula, for every norm (see Ziętak [26]). Miao and Ben-Israel [15] have given another expression for the l_p-approximate generalized inverses of $A \in \mathscr{R}_n^{(n+1) \times n}$. The approach of Miao and Ben-Israel was inspired by the paper of Berg [4] (see also Ben-Israel [2]). Their formulae can be easily extended to the case of more general norms using the results presented in this paper (see Ziętak [27]).

ACKNOWLEDGMENT

The author thanks the referee for his remarks which helped to improve the presentation.

REFERENCES

1. F. L. Bauer, J. Stoer, and C. Witzgall, Absolute and monotonic norms, *Numer. Math.* **3** (1961), 257–264.
2. A. Ben-Israel, A volume associated with $m \times n$ matrices, *Linear Algebra Appl.* **167** (1992), 87–111.
3. A. Ben-Tall and M. Teboulle, A geometric property of the least squares solution of linear equations, *Linear Algebra Appl.* **139** (1990), 165–170.
4. L. Berg, Three results in connection with inverse matrices, *Linear Algebra Appl.* **84** (1986), 63–77.
5. E. W. Cheney, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
6. J. Descloux, Approximation in L^p and Chebyshev approximation, *J. Soc. Ind. Appl. Math.* **11** (1963), 1017–1026.

7. A. Forsgren, "On Linear Least-Squares Problem with Diagonally Dominant Weight Matrices," Report TRITA/MAT-95-0S2, Department of Mathematics, Royal Institute of Technology, Stockholm, 1995.
8. D. Gries, Characterizations of certain classes of norms, *Numer. Math.* **10** (1967), 30–41.
9. D. Gries and J. Stoer, Some results on fields of values of a matrix, *SIAM J. Numer. Anal.* **4** (1967), 283–300.
10. M. Hanke and M. Neumann, The geometry of the set of scaled projections, *Linear Algebra Appl.* **190** (1993), 137–148.
11. D. P. O'Leary, On bounds for scaled projections and pseudoinverses, *Linear Algebra Appl.* **132** (1990), 115–117.
12. M. L. Levitan and R. Y. S. Lynn, An overdetermined linear system, *J. Approx. Theory* **18** (1976), 264–277.
13. G. Loizou, Diagonal transformations, *Appl. Math.* **20** (1990), 245–254.
14. M. Meicler, Chebyshev solution of an inconsistent system of $n+1$ linear equations in n unknowns in terms of its least squares solution, *SIAM Rev.* **10** (1968), 373–375.
15. J. Miao and A. Ben-Israel, On l_p-approximate solutions of linear equations, *Linear Algebra Appl.* **199** (1994), 305–327.
16. J. Miao and A. Ben-Israel, The geometry of basic, approximate, and minimum norm solutions, *Linear Algebra Appl.* **216** (1995), 25–42.
17. J. M. Ortega and W. C. Rheinbold, "Iterative Solution of Nonlinear Equations in Several Variables," Academic Press, San Diego, 1970.
18. J. Rice, "The Approximation of Functions," Vol. II, Addison–Wesley, Reading, MA, 1969.
19. E. Marques de Sá and M.-J. Sodupe, Characterizations of *orthant-monotonic norms, *Linear Algebra Appl.* **193** (1993), 1–9.
20. V. P. Sreedharan, Solutions of overdetermined linear equations which minimize error in an abstract norm, *Numer. Math.* **13** (1969), 146–151.
21. V. P. Sreedharan, Least squares algorithms for finding solutions of overdetermined linear equations which minimize error in an abstract norm, *Numer. Math.* **17** (1971), 387–401.
22. V. P. Sreedharan, Least squares algorithms for finding solutions of overdetermined systems of linear equations which minimize error in a smooth strictly convex norm, *J. Approx. Theory* **8** (1973), 46–61.
23. G. W. Stewart, On scaled projections and pseudoinverses, *Linear Algebra Appl.* **112** (1989), 189–193.
24. J. Stoer and C. Witzgall, Transformation by diagonal matrices in a normed space, *Numer. Math.* **4** (1962), 158–171.
25. G. A. Watson, "Approximation Theory and Numerical Methods," Wiley, Chichester, 1980.
26. K. Ziętak, Properties of the approximate generalized inverses of a class of matrices, *Linear Algebra Appl.* **92** (1987), 259–272.
27. K. Ziętak, Approximate and minimum generalized inverses of a matrix, in preparation.