Lin. Algeb. and its Appl. **216**(1995) 25–42

The geometry of basic, approximate, and minimum norm solutions of linear equations

Jianming Miao* and Adi Ben-Israel †

June 30, 1992

Abstract

The **basic solutions** of the linear equation

$$A\mathbf{x} = \mathbf{b} \ ,$$

are the solutions of subsystems corresponding to maximal nonsingular submatrices of A. The **convex hull** of the basic solutions is denoted by $\mathcal{C} = \mathcal{C}(A, \mathbf{b})$. The **residual** $\mathbf{r}(\mathbf{x})$ of a vector $\mathbf{x}$ is $\mathbf{r} := A\mathbf{x} - \mathbf{b}$. Given $1 \le p \le \infty$, the **ℓ_p-approximate solutions** of $A\mathbf{x} = \mathbf{b}$, denoted $\mathbf{x}^{\{p\}}$, are minimizers of $\|\mathbf{r}(\mathbf{x})\|_p$. Given $M \in \mathcal{D}_m$, the set of positive diagonal $m \times m$ matrices, the solutions of

$$\min_{\mathbf{x}} \|M\,(A\mathbf{x} - \mathbf{b})\|_p \ ,$$

are called **scaled ℓ_p-approximate solutions**. For $1 \le p_1, p_2 \le \infty$, the **minimum ℓ_{p_2}-norm ℓ_{p_1}-approximate solutions** are denoted $\mathbf{x}^{\{p_1\}}_{\{p_2\}}$. Main results:

(a) The set of scaled ℓ_p-approximate solutions, with M ranging over $\mathcal{D}_m$, is the same for all $1 < p < \infty$.
(b) If $A \in \mathbb{R}^{m\times n}_m$, $\mathcal{C}$ contains all [some] minimum ℓ_p-norm solutions, for $1 \le p < \infty$ [$p = \infty$].
(c) For general A, $1 \le p_1, p_2 < \infty$, the set $\mathcal{C}$ contains all $\mathbf{x}^{\{p_1\}}_{\{p_2\}}$.

Key words: Linear equations. ℓ_p-approximate solutions. Minimum ℓ_p-norm solutions. Least squares solutions.

---

*Supported by Grant NSF-STC91-19999.
†RUTCOR-Rutgers Center for Operations Research, Rutgers University, P.O. Box 5062, New Brunswick, NJ 08903-5062, U.S.A.

0 Notation

0.1 The set of increasing sequences of k elements from $\{1, \cdots, m\}$ is denoted by

$$Q_{k,m} = \{I = \{i_1, \cdots, i_k\} : 1 \le i_1 < i_2 < \cdots < i_k \le m\} .$$

Given a subset $I \in Q_{k,m}$, its **complement** in $\{1, \ldots, m\}$ is denoted I^c.

0.2 For $A \in \mathbb{R}_r^{m \times n}$, $r > 0$, denote the **index sets**

$$\begin{aligned} \mathcal{I}(A) &= \{I \in Q_{r,m} : \text{ rank } A_{I*} = r\} , & (0.1)\\ \mathcal{J}(A) &= \{J \in Q_{r,n} : \text{ rank } A_{*J} = r\} , & (0.2)\\ \mathcal{N}(A) &= \{(I, J) \in Q_{r,m} \times Q_{r,n} : \text{ rank } A_{IJ} = r\} , & (0.3) \end{aligned}$$

of **maximal sets of linearly independent rows** and **columns**, and of **maximal nonsingular submatrices**, respectively. The index sets $\mathcal{I}(A)$, $\mathcal{J}(A)$ and $\mathcal{N}(A)$ are abbreviated here by $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{N}$ respectively. We have

$$\mathcal{N} = \mathcal{I} \times \mathcal{J} , \quad \text{see e.g. [2]} . \qquad (0.4)$$

0.3 The **basic solutions** of the linear equation

$$A\,\mathbf{x} = \mathbf{b} ,$$

are the solutions of subsystems corresponding to maximal nonsingular submatrices of A. The basic solutions are, for

$$\begin{aligned} A \text{ of full column-rank:} &\quad \{A_{I*}^{-1}\,\mathbf{b}_I \;:\; I \in \mathcal{I}\} , & (0.5)\\ A \text{ of full row-rank:} &\quad \{\widehat{A_{*J}^{-1}}\,\mathbf{b} \;:\; J \in \mathcal{J}\} , & (0.6)\\ \text{general } A: &\quad \{\widehat{A_{IJ}^{-1}}\,\mathbf{b}_I \;:\; (I, J) \in \mathcal{N}\} , & (0.7) \end{aligned}$$

where $\mathbf{b}_I$ is the I^{th} subvector of $\mathbf{b}$, and $\frown$ denotes a vector padded by zeros.

The **convex hull of basic solutions** of the given equation $A\,\mathbf{x} = \mathbf{b}$ is denoted by $\mathcal{C} = \mathcal{C}(A, \mathbf{b})$.

0.4 The **set of minimizers** [**maximizers**] of a function f is denoted by arg min f [arg max f].

0.5 The **Hadamard product** $\mathbf{u} \circ \mathbf{v}$ of two vectors $\mathbf{u} = (u_j)$ and $\mathbf{v} = (v_j)$ is the vector

$$\mathbf{u} \circ \mathbf{v} := (u_j\, v_j) . \qquad (0.8)$$

0.6 Let $\mathcal{D}_m$ be the set of all $m \times m$ positive diagonal matrices.

0.7 Inequalities between vectors, such as $\mathbf{x} \le \mathbf{y}$, are interpreted componentwise. $\mathbf{x} \lneq \mathbf{y}$ would be $\mathbf{x} \le \mathbf{y}\,, \mathbf{x} \ne \mathbf{y}$.

1 Introduction

Given $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$, consider the linear equation

$$A\,\mathbf{x} = \mathbf{b} , \qquad (1.1)$$

and for any $\mathbf{x} \in \mathbb{R}^n$, the **residual vector**

$$\mathbf{r} = \mathbf{r}(\mathbf{x}) := A\mathbf{x} - \mathbf{b} . \qquad (1.2)$$

If (1.1) is inconsistent, we often settle for an approximate solution minimizing a norm of the residual. Using the family of ℓ_p**-norms**, $1 \le p \le \infty$, defined for $\mathbf{u} = (u_j) \in \mathbb{R}^m$ by

$$\|\mathbf{u}\|_p \;:=\; \begin{cases} \left(\sum_{j=1}^m |u_j|^p \right)^{\frac{1}{p}} & , \; 1 \le p < \infty ,\\ \max_{1 \le j \le m} |u_j| & , \; p = \infty , \end{cases} \qquad (1.3)$$

a ℓ_p**-approximate solution** of (1.1) is a solution of the minimization problem

$$\min \, \{\|A\,\mathbf{x} - \mathbf{b}\|_p \;:\; \mathbf{x} \in \mathbb{R}^n\} . \qquad (1.4)$$

In particular, the ℓ_2-norm is the **Euclidean norm**

$$\|\mathbf{u}\|_2 \;:=\; \sqrt{\sum_{j=1}^m u_j^2} , \qquad (1.5)$$

and the ℓ_2-approximate solutions are the **least squares solutions**.

For A of full column-rank, Berg [5] proved the least squares solution is in

$$\mathcal{C} := \text{conv}\{\; A_{I*}^{-1}\,\mathbf{b}_I \;:\; I \in \mathcal{I} \;\} ,$$

the convex hull of basic solutions (0.5)[1]. Ben-Tal and Teboulle [4] extended the results to isotone functions, of which ℓ_p-norms can be considered a special case. A continuous function $f : \mathbb{R}_+^m \;\rightarrow\; \mathbb{R}$ is called **isotone** if

$$f(\mathbf{x}) \le f(\mathbf{y}) , \quad \text{whenever} \quad 0 \le \mathbf{x} \le \mathbf{y} , \qquad (1.6)$$

and **strictly isotone** if in addition,

$$0 \le \mathbf{x} \le \mathbf{y}, \quad f(\mathbf{x}) = f(\mathbf{y}) \quad \Longrightarrow \quad \mathbf{x} = \mathbf{y} . \qquad (1.7)$$

For any $1 \le p \le \infty$ [$1 \le p < \infty$], the l_p-norm $\|\mathbf{x}\|_p$ is [strictly] isotone function of the vector $|\mathbf{x}|$ of absolute values,

$$|\mathbf{x}| = (|x_1|, \;\cdots, \;|x_n|)^T . \qquad (1.8)$$

Lemma 1.1 ([4]) Let $A \in \mathbb{R}_n^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and let $f : \mathbb{R}_+^m \to \mathbb{R}$ be isotone. Then the problem

$$\min_{\mathbf{x}} \; f(|A\mathbf{x} - \mathbf{b}|) , \qquad (1.9)$$

has a solution in $\mathcal{C}$. Moreover, if f is strictly isotone, then every solution of (1.9) lies in $\mathcal{C}$. □

[1] This is important for establishing convergence of certain iterative methods since the set $\mathcal{C}$ is compact.

These results are extended here along the following lines:

• Geometrical properties of scaled ℓ_p-approximate solutions are studied in Section 2 for A of full column-rank. We show that for $1 < p < \infty$, the set of scaled ℓ_p-approximate solutions is the same as the set of scaled least squares solutions.

• In Section 3 we consider the problem

$$\min_{\mathbf{x}} \{f(|\mathbf{x}|) \ : \ A\mathbf{x} = \mathbf{b}\} ,$$

where A is a matrix of full row-rank and f is isotone. We show that there is a solution in $\mathcal{C}$. Moreover, if f is strictly isotone then every solution lies in $\mathcal{C}$.

• In § 4 we consider the problem

$$\min_{\mathbf{x}} \{f_2(|\mathbf{x}|) \ : \ \mathbf{x} \in \arg \min_{\mathbf{x}} \ f_1(|A\mathbf{x} - \mathbf{b}|)\} , \tag{1.10}$$

where $A \in \mathbb{R}_r^{m\times n}$. For f_2 isotone and f_1 strictly isotone, $\mathcal{C}$ contains a solution of (1.10). If also f_2 is strictly isotone, then every solution of (1.10) lies in $\mathcal{C}$.

2 A is of full column-rank

Throughout this section let $A \in \mathbb{R}_n^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. The convex hull of the basic solutions (0.5) is

$$\mathcal{C} \ := \ \text{conv}\{A_{I*}^{-1}\,\mathbf{b}_I \ : \ I \in \mathcal{I}\} . \tag{2.1}$$

For any $1 \le p \le \infty$ and $M \in \mathcal{D}_m$ (see § 0.6), consider the problem

$$\min_{\mathbf{x}} \ \|M\,(A\mathbf{x} - \mathbf{b})\|_p , \tag{2.2}$$

whose solution is unique for $1 < p < \infty$. The solutions are called **scaled ℓ_p-approximate solutions**. For $p = 2$, $D \in \mathcal{D}_m$, the **scaled least squares solution** of $A\,\mathbf{x} = \mathbf{b}$ is the solution of

$$\min_{\mathbf{x}} \ \|D^{\frac{1}{2}}\,(A\mathbf{x} - \mathbf{b})\|_2 , \tag{2.3}$$

given by

$$\mathbf{x} = (A^TDA)^{-1}A^TD\mathbf{b} , \quad \text{see e.g. [3]} . \tag{2.4}$$

Let the set of scaled ℓ_p-approximate solutions be

$$\mathcal{X}^{\{p\}} \ := \ \bigcup_{M\in\mathcal{D}_m} \left\{ \arg \min_{\mathbf{x}} \|M\,(A\mathbf{x} - \mathbf{b})\|_p \right\} , \tag{2.5}$$

and for $p = 2$, by (2.4),

$$\mathcal{X}^{\{2\}} \ = \ \{ \ (A^TDA)^{-1}A^TD\mathbf{b} \ : \ D \in \mathcal{D}_m \ \} . \tag{2.6}$$

For $1 < p < \infty$, each arg min in (2.5) is a singleton.

Theorem 2.1 Let $A \in \mathbb{R}_n^{m\times n}$, $1 < p < \infty$. Then $\mathcal{X}^{\{p\}} = \mathcal{X}^{\{2\}}$.

Proof. The result is trivially true if $\mathbf{b} \in R(A)$, the **range** of A.

Let $\mathbf{b} \notin R(A)$. The function $f(\mathbf{x}) := \|M\,(A\mathbf{x} - \mathbf{b})\|_p$ is convex and differentiable, and a point $\mathbf{x}^*$ is the optimal solution of (2.2) if and only if

$$\nabla f(\mathbf{x}^*) = \mathbf{0} .$$

$\underline{\mathcal{X}^{\{p\}} \subset \mathcal{X}^{\{2\}}}$: Let $\mathbf{x}^*$ be the solution of (2.2), and define

$$\overline{A} \ := \ MA, \ \ \overline{\mathbf{b}} \ := \ M\,\mathbf{b}, \ \ \overline{\mathbf{r}}(\mathbf{x}^*) \ := \ \overline{A}\mathbf{x}^* - \overline{\mathbf{b}} . \tag{2.7}$$

Then (see e.g. [8, Theorem 2.1])

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \quad \Longleftrightarrow \quad \overline{A}^T \left(\overline{\mathbf{r}}(\mathbf{x}^*) \circ |\overline{\mathbf{r}}(\mathbf{x}^*)|^{p-2}\right) = \mathbf{0} , \tag{2.8}$$

Let the diagonal matrix $\overline{M} = \text{diag}\,(\overline{m}_j)$ be defined by

$$\overline{m}_j \ := \ \begin{cases} |\overline{r}_j(\mathbf{x}^*)|^{p-2} & , \text{ if } \overline{r}_j(\mathbf{x}^*) \neq 0 , \\ 1 & , \text{ otherwise } . \end{cases}$$

Then (2.8) gives

$$\overline{A}^T\overline{M}(\overline{A}\mathbf{x}^* - \overline{\mathbf{b}}) \ = \ \mathbf{0}.$$

Therefore

$$\mathbf{x}^* = (A^TDA)^{-1}A^TD\mathbf{b} \ \in \ \mathcal{X}^{\{2\}} , \quad \text{where } D \ := \ M\overline{M}M .$$

$\underline{\mathcal{X}^{\{2\}} \subset \mathcal{X}^{\{p\}}}$: Let $\mathbf{x}^*$ be any scaled least squares solution, i.e. $\mathbf{x}^*$ satisfies

$$A^TD\,(A\mathbf{x}^* - \mathbf{b}) = \mathbf{0}, \ \text{ for some } D \in \mathcal{D}_m . \tag{2.9}$$

Let $\mathbf{r}(\mathbf{x}^*) = A\mathbf{x}^* - \mathbf{b}$, and define the matrix $M = \text{diag}\,(m_j)$ by

$$m_j \ := \ \begin{cases} \sqrt[p]{\dfrac{d_j}{|r_j(\mathbf{x}^*)|^{p-2}}} & , \text{ if } r_j(\mathbf{x}^*) \neq 0 , \\ 1 & , \text{ otherwise } . \end{cases}$$

Then (2.9) gives

$$\overline{A}^T \left(\overline{\mathbf{r}}(\mathbf{x}^*) \circ |\overline{\mathbf{r}}(\mathbf{x}^*)|^{p-2}\right) \ = \ \mathbf{0} , \tag{2.10}$$

where

$$\overline{A} \ := \ MA, \ \ \overline{\mathbf{b}} \ := \ M\,\mathbf{b}, \ \ \overline{\mathbf{r}}(\mathbf{x}^*) \ := \ \overline{A}\mathbf{x}^* - \overline{\mathbf{b}} .$$

By (2.8), $\mathbf{x}^*$ is the solution of (2.2). □

Remark 2.1 We prove now that $\mathcal{X}^{\{2\}} \subset \mathcal{X}^{\{p\}}$ for $p = 1$ and $p = \infty$ by immitating the proof of $\mathcal{X}^{\{2\}} \subset \mathcal{X}^{\{p\}}$ in Theorem 2.1. As there, let $\mathbf{x}^*$ be any scaled least squares solution.

$\underline{\mathcal{X}^{\{2\}} \subset \mathcal{X}^{\{1\}}}$: For $p = 1$, (2.10) becomes

$$\overline{A}^T\,(\text{sgn}\,\overline{\mathbf{r}}(\mathbf{x}^*)) = \mathbf{0},$$

where $\operatorname{sgn}\overline{\mathbf{r}}(\mathbf{x}^*) = (\operatorname{sgn}\overline{r}_i(\mathbf{x}^*))$ the signum vector. Using [8, Theorem 2.1] we conclude that $\mathbf{x}^*$ is a solution of (2.2) for $p = 1$.

$\underline{\mathcal{X}^{\{2\}} \subset \mathcal{X}^{\{\infty\}}}$: Let $p = \infty$, and define the matrix $M = \operatorname{diag}(m_j)$ by

$$m_j := \begin{cases} \dfrac{\sum_{i=1}^{m} d_i |r_i(\mathbf{x}^*)|^2}{|r_j(\mathbf{x}^*)|}, & \text{if } r_j(\mathbf{x}^*) \neq 0 , \\ 1 , & \text{otherwise} . \end{cases}$$

Then

$$M^{-1} D\,\mathbf{r}(\mathbf{x}^*) \in$$

$$\operatorname{conv}\{\, (\operatorname{sgn}\overline{r}_j(\mathbf{x}^*))\,\mathbf{e}_j \;:\; |\overline{r}_j(\mathbf{x}^*)| = \|\overline{\mathbf{r}}(\mathbf{x}^*)\|_\infty \,\} \cap N(\overline{A}^T) ,$$

where $N(\cdot)$ denotes null space. By [8, Theorem 2.1], $\mathbf{x}^*$ is a solution of (2.2) for $p = \infty$. $\square$

If the norm $\|\cdot\|$ is not isotone (isotone norms are also called **monotone**), then the solutions of $\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|$ may lie outside $\mathcal{C}$.

Example 2.1 Let

$$W = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} .$$

Then the norm $\|\mathbf{x}\|_W := \|W^{\frac{1}{2}}\mathbf{x}\|_2$ is not isotone. For

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} ,$$

the basic solutions are

$$\mathbf{x}_1 = -1, \quad \mathbf{x}_2 = 1,$$

and their convex hull is the interval

$$\mathcal{C} = [-1,\ 1] .$$

Finally, the solution of $\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_W$ is

$$\begin{aligned} \mathbf{x} &= (A^T W A)^{-1} A^T W \mathbf{b} \\ &= 2 \notin \mathcal{C}. \end{aligned}$$

The following example shows that in general $\mathcal{X}^{\{2\}} \neq \mathcal{X}^{\{\infty\}}$.

Example 2.2 (Based on [6, Example 5.2]). Let

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -2 & 2 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ -2 \\ 2 \end{pmatrix} .$$

The left plot of Figure 1 shows $\mathcal{X}^{\{2\}}$, which consists of the interiors of the two shaded triangles and their common point $\mathbf{x} = (2,\ 0)$. The ℓ_∞-approximate solutions are on the line segment X. Finally, the set $\mathcal{X}^{\{\infty\}}$ consists of all points between the two lines $L_1,\ L_2$ (excluding $L_1,\ L_2$).

Ben-Tal and Teboulle proved $\mathcal{X}^{\{2\}} \subset \mathcal{C}$. Recently Hanke and Neumann [6] showed $\mathcal{X}^{\{2\}}$ to be a union of finitely many polytopes, in general not convex, and $c\ell\,\mathcal{X}^{\{2\}} \subset \mathcal{C}$, where $c\ell$ denotes **closure**. The results of [6] and Theorem 2.1 imply that not all vectors in $\mathcal{C}$ are scaled ℓ_p-approximate solutions for $1 < p < \infty$. The next example shows not all vectors in $\mathcal{C}$ are solutions of $\min_{\mathbf{x}} f(|A\mathbf{x} - \mathbf{b}|)$ for some strictly isotone function f.

Example 2.3 (Based on [6, Example 5.1]). Let

$$A = \begin{pmatrix} 2 & -2 \\ 1 & 0 \\ 2 & 8 \\ 2 & -6 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ 0 \\ 3 \\ 3 \end{pmatrix} .$$

The right plot of Figure 1 shows the convex hull $\mathcal{C}$ of basic solutions (the triangle bounded by thick lines), and the set $c\ell\,\mathcal{X}^{\{2\}}$ (the shaded region).

Consider the points $\mathbf{x} = \begin{pmatrix} 1 \\ \frac{3}{8} \end{pmatrix} \in \mathcal{C} \setminus c\ell\,\mathcal{X}^{\{2\}}$ and $\mathbf{y} = \begin{pmatrix} \frac{2}{3} \\ 0 \end{pmatrix} \in \mathcal{X}^{\{2\}}$. Then

$$|A\mathbf{x} - \mathbf{b}| = \begin{pmatrix} \frac{19}{4} \\ 1 \\ 2 \\ \frac{13}{4} \end{pmatrix} > |A\mathbf{y} - \mathbf{b}| = \begin{pmatrix} \frac{14}{3} \\ \frac{2}{3} \\ \frac{5}{3} \\ \frac{5}{3} \end{pmatrix} ,$$

which implies

$$f(|A\mathbf{x} - \mathbf{b}|) > f(|A\mathbf{y} - \mathbf{b}|) ,$$

for any strictly isotone function f, showing that the point $\mathbf{x}$ is not a solution of $\min_{\mathbf{x}} f(|A\mathbf{x} - \mathbf{b}|)$.

Let $\mathcal{F}_m$ be the set of all strictly isotone functions on $\mathbb{R}^m$, and let

$$\mathcal{X}^{\{F\}} := \bigcup_{f \in \mathcal{F}_m} \left\{ \mathbf{x} \;:\; \mathbf{x} \in \arg\min_{\mathbf{x}} f(|A\mathbf{x} - \mathbf{b}|) \right\} . \tag{2.11}$$

The question,

$$c\ell\,\mathcal{X}^{\{2\}} \stackrel{?}{=} c\ell\,\mathcal{X}^{\{F\}} ,$$

suggested by Example 2.3, is answered in the affirmative, in Theorem 2.4. First we need some other results.

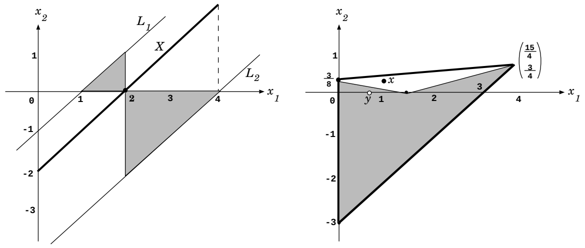

Figure 1: Illustration of Examples 2.2 and 2.3
.

Let $\mathcal{S}$ be a polytope in $\mathbb{R}^m$,

$$\mathcal{S} = \{ \ \mathbf{x} = \sum_{i=1}^{k} \lambda_i \, \mathbf{x}^i \ : \ \sum_{i=1}^{k} \lambda_i \ = 1, \ \lambda_i \geq 0, \ i = 1, \ \ldots, \ k \ \} \ , \tag{2.12}$$

such that $\mathbf{0} \notin \mathcal{S}$. For any $D \in \mathcal{D}_m$, denote

$$\mathbf{x}_D = \arg \min_{\mathbf{x} \in \mathcal{S}} \| D \, \mathbf{x} \|_2 \ . \tag{2.13}$$

We define

$$\mathcal{P} \ = \ \{ \ \mathbf{x}_D \ : \ D \in \mathcal{D}_m \ \}, \tag{2.14}$$

$$\mathcal{A} \ = \ \{ \ \mathbf{x} \in \mathcal{S} \ : \nexists \ \mathbf{y} \in \mathcal{S} \text{ such that } |\mathbf{y}| \ \lneqq \ |\mathbf{x}| \ \}. \tag{2.15}$$

Lemma 2.1 Let $\mathbf{x} \in \mathbb{R}^m$. Then

$$\mathbf{x} \in \mathcal{P} \iff Z \, \mathbf{p} \ \lneqq \ \mathbf{0}, \ \mathbf{p} \geq \mathbf{0} \ \text{ has no solution,} \tag{2.16}$$

where $Z = \left(\mathbf{z}^1, \ \mathbf{z}^2, \ \cdots, \ \mathbf{z}^k \right)$ is the matrix with columns

$$\mathbf{z}^i = \mathbf{x} \circ (\mathbf{x}^i - \mathbf{x}), \ i = 1, \ \ldots, \ k. \tag{2.17}$$

Proof. $\mathbf{x} \in \mathcal{P} \iff$

$$\begin{array}{lll}
\iff & \exists \ D \in \mathcal{D}_m, \ \ < D \, (\mathbf{y} - \mathbf{x}), \ D \, \mathbf{x} > \ \geq \ 0, & \forall \ \mathbf{y} \in \mathcal{S}, \\
 & & [1, \text{ p. } 41], \\
\iff & \exists \ D \in \mathcal{D}_m, \ \ \mathbf{x}^T D^2 \, (\mathbf{y} - \mathbf{x}) \ \geq \ 0, & \forall \ \mathbf{y} \in \mathcal{S}, \\
\iff & \exists \ D \in \mathcal{D}_m, \ \ \mathbf{x}^T D^2 \, (\mathbf{x}^i - \mathbf{x}) \ \geq \ 0, & i = 1, \ \ldots, \ k, \\
\iff & Z^T \mathbf{d} \geq \mathbf{0}, \quad \mathbf{d} > \mathbf{0} \quad \text{has a solution.} &
\end{array}$$

By a theorem of alternatives [7, p. 29]

$$\mathbf{x} \in \mathcal{P} \iff Z \, \mathbf{p} \ \lneqq \ \mathbf{0}, \ \mathbf{p} \geq \mathbf{0} \ \text{ has no solution.} \qquad \square$$

Theorem 2.2 $\mathcal{P} \subset \mathcal{A}$.

Proof. For any $\mathbf{x} \in \mathcal{S} \backslash \mathcal{A}$, there is $\mathbf{y} \in \mathcal{S}$ such that $|\mathbf{y}| \ \lneqq \ |\mathbf{x}|$. Therefore

$$\| D \, \mathbf{y} \|_2 < \| D \, \mathbf{x} \|_2 \ ,$$

for any $D \in \mathcal{D}_m$, which implies $\mathbf{x} \in \mathcal{S} \setminus \mathcal{P}$. □

Theorem 2.3 $\mathcal{A} \subset c\ell \, \mathcal{P}$.

Proof.
Case 1. $\quad \mathbf{x} = (x_i) \in \mathcal{A} \ , \quad x_i \neq 0, \ i = 1, \ \ldots, \ m \ .$
We show that $\mathbf{x} \in \mathcal{P}$. If not, then by Lemma 2.1

$$Z \, \mathbf{p} \ \lneqq \ \mathbf{0}, \quad \mathbf{p} \geq \mathbf{0}, \tag{2.18}$$

has a solution $\mathbf{p}$. Let

$$\mathbf{y} \ := \ \sum_{i=1}^{k} \lambda_i \, \mathbf{x}^i \ \in \ \mathcal{S},$$

with

$$\lambda_j \ := \ \frac{p_j}{\sum_{i=1}^{k} p_i}, \quad j = 1, \ 2, \ \ldots, \ k \ .$$

Then (2.18) gives

$$\mathbf{x} \circ (\mathbf{y} - \mathbf{x}) \ \lneqq \ \mathbf{0}. \tag{2.19}$$

For sufficiently small $\lambda > 0$, the vector

$$\mathbf{z} \ := \ \lambda \, \mathbf{y} + (1 - \lambda) \, \mathbf{x} \ \in \ \mathcal{S}.$$

Then it follows from (2.19) that

$$|\mathbf{z}| \ \lneqq \ |\mathbf{x}| \ , \quad \text{contradicting} \quad \mathbf{x} \in \mathcal{A} \ .$$

Case 2. $\quad \mathbf{x} = (x_i) \in \mathcal{A} \ , \quad I^c = \{ i \ : \ x_i = 0 \} \neq \emptyset.$
Without loss of generality let $\mathbf{x} = \begin{pmatrix} \mathbf{x}_I \\ \mathbf{0} \end{pmatrix}$. We define

$$\mathcal{S}_I \ := \ \{ \ \mathbf{y}_I \ : \ \begin{pmatrix} \mathbf{y}_I \\ \mathbf{0} \end{pmatrix} \in \mathcal{S} \ \}, \tag{2.20}$$

$$\mathcal{A}_I \ := \ \{ \ \mathbf{x}_I \in \mathcal{S}_I \ : \nexists \ \mathbf{y}_I \in \mathcal{S}_I \text{ such that } |\mathbf{y}_I| \ \lneqq \ |\mathbf{x}_I| \ \} \tag{2.21}$$

Then $\mathcal{S}_I$ is a polytope and $\mathbf{x}_I \in \mathcal{A}_I$. By case 1, there is a positive diagonal matrix D_I such that

$$\mathbf{x}_I = \arg \min_{\mathbf{y}_I \in \mathcal{S}_I} \| D_I \, \mathbf{y}_I \|_2 \ . \tag{2.22}$$

Let $D_n = \begin{pmatrix} \frac{1}{n} D_I & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} \in \mathcal{D}_m$, and let $\mathbf{x}_n := \mathbf{x}_{D_n}$. Then by the definition (2.13)

$$\left\| \begin{pmatrix} \frac{1}{n} D_I \, (\mathbf{x}_n)_I \\ (\mathbf{x}_n)_{I^c} \end{pmatrix} \right\|_2 \ \leq \ \left\| \begin{pmatrix} \frac{1}{n} D_I \, \mathbf{x}_I \\ \mathbf{0} \end{pmatrix} \right\|_2 \ . \tag{2.23}$$

Since $\mathcal{S}$ is bounded, the sequence $\{\mathbf{x}_n\}$ has a convergent subsequence. Without loss of generality, let $\mathbf{x}_n \longrightarrow \overline{\mathbf{x}} \in c\ell \, \mathcal{P}$. Then it follows from (2.23) that

$$\overline{\mathbf{x}}_{J^c} = \mathbf{0},$$

and

$$\| D_I \, \overline{\mathbf{x}}_I \|_2 \ \leq \ \| D_I \, \mathbf{x}_I \|_2 \ .$$

By the uniqueness of $\mathbf{x}_I$ in (2.22), we have $\mathbf{x} = \overline{\mathbf{x}} \in c\ell \, \mathcal{P}$. □

Theorem 2.4 $c\ell\,\mathcal{X}^{\{2\}} = c\ell\,\mathcal{X}^{\{F\}}$.

Proof. $c\ell\,\mathcal{X}^{\{2\}} \subset c\ell\,\mathcal{X}^{\{F\}}$ is obviously true. We prove $c\ell\,\mathcal{X}^{\{F\}} \subset c\ell\,\mathcal{X}^{\{2\}}$ by showing

$\underline{\mathcal{X}^{\{F\}} \subset c\ell\,\mathcal{X}^{\{2\}}}$: Let $\mathcal{S}$ be the polytope defined by

$$\mathcal{S} = \{ \, \mathbf{r}(\mathbf{x}) = A\mathbf{x} - \mathbf{b} \;:\; \mathbf{x} \in \mathcal{C} \, \}.$$

Define $\mathcal{P}$, $\mathcal{A}$ as before, and let $\mathbf{x} \notin c\ell\,\mathcal{X}^{\{2\}}$. Then $\mathbf{r}(\mathbf{x}) \notin c\ell\,\mathcal{P}$. By Theorem 2.3, $\mathbf{r}(\mathbf{x}) \notin \mathcal{A}$. Therefore there is $\mathbf{y} \in \mathcal{C}$ such that

$$|A\mathbf{y} - \mathbf{b}| \;\lneqq\; |A\mathbf{x} - \mathbf{b}|,$$

which implies

$$f(|A\mathbf{y} - \mathbf{b}|) < f(|A\mathbf{x} - \mathbf{b}|)$$

for any $f \in \mathcal{F}_m$. Therefore $\mathbf{x} \notin \mathcal{X}^{\{F\}}$, proving that $\mathcal{X}^{\{F\}} \subset c\ell\,\mathcal{X}^{\{2\}}$. □

Theorem 2.5 $c\ell\,\mathcal{X}^{\{1\}} = c\ell\,\mathcal{X}^{\{2\}}$.

Proof. Follows from Remark 2.1 and Theorem 2.4. □

3 A is of full row-rank

Throughout this section let $A \in \mathbb{R}^{m\times n}_m$ and $\mathbf{b} \in \mathbb{R}^m$. The convex hull of the basic solutions (0.6) is

$$\mathcal{C} \;:=\; \mathrm{conv}\{\widehat{A^{-1}_{*J}\mathbf{b}} \;:\; J \in \mathcal{J}\} \,, \tag{3.1}$$

where $\widehat{A^{-1}_{*J}}\,\mathbf{b}$ has $A^{-1}_{*J}\,\mathbf{b}$ in position J, zeros elsewhere. For any $1 \le p \le \infty$ and $N \in \mathcal{D}_n$, consider the problem

$$\min_{\mathbf{x}} \, \{ \, \|N^{-1}\,\mathbf{x}\|_p : \; A\mathbf{x} = \mathbf{b} \, \} \,, \tag{3.2}$$

and its solutions, called **scaled minimum ℓ_p-norm solutions**, which are unique for $1 < p < \infty$.

If $N = I$, these solutions are simply called **minimum ℓ_p-norm solutions**.

For $p = 2$ and any $D \in \mathcal{D}_n$, the **scaled minimum ℓ_2-norm solution** of

$$\min_{\mathbf{x}} \, \{ \, \|D^{-\frac{1}{2}}\,\mathbf{x}\|_2 : \; A\mathbf{x} = \mathbf{b} \, \} \,, \tag{3.3}$$

is easily computed (see, e.g. [3])

$$\mathbf{x} = DA^T(ADA^T)^{-1}\mathbf{b} \,. \tag{3.4}$$

Let the set of scaled minimum ℓ_p-norm solutions be

$$\mathcal{X}_{\{p\}} \;:=\; \bigcup_{N\in\mathcal{D}_n} \left\{ \, \mathbf{x} \;:\; \mathbf{x} \in \arg\min_{\mathbf{x}} \{\|N^{-1}\,\mathbf{x}\|_p : \; A\mathbf{x} = \mathbf{b}\} \, \right\} . \tag{3.5}$$

Then (3.4) gives

$$\mathcal{X}_{\{2\}} \;=\; \{ \, DA^T(ADA^T)^{-1}\mathbf{b} \;:\; D \in \mathcal{D}_n \, \}. \tag{3.6}$$

Lemma 3.1 Let $A \in \mathbb{R}^{m\times n}_m$. Then $\mathcal{X}_{\{2\}} \subset \mathcal{C}$.

Proof. Let $\mathbf{x}$ be the solution of (3.3), $\mathbf{y} := D^{-\frac{1}{2}}\mathbf{x}$, $B := AD^{\frac{1}{2}}$. Then $\mathbf{y}$ is the minimum ℓ_2-norm solution of $B\mathbf{y} = \mathbf{b}$, and by [2], a convex combination of basic solutions,

$$\mathbf{y} = \sum_{J\in\mathcal{J}} \gamma_J \, \widehat{B^{-1}_{*J}}\mathbf{b} \,.$$

Therefore

$$\begin{aligned} \mathbf{x} \;&=\; D^{\frac{1}{2}}\mathbf{y} \,, \\ &=\; \sum_{J\in\mathcal{J}} \gamma_J \, \widehat{A^{-1}_{*J}}\mathbf{b} \;\in \mathcal{C} \,. \end{aligned}$$

□

Theorem 3.1 Let $A \in \mathbb{R}^{m\times n}_m$, $1 < p < \infty$. Then $\mathcal{X}_{\{p\}} = \mathcal{X}_{\{2\}}$.

Proof. A suitable Lagrangian function is

$$\mathcal{L}(\mathbf{x},\; \mathbf{u}) \;:=\; \|N^{-1}\mathbf{x}\|^p_p - p\,\mathbf{u}^T(A\mathbf{x} - \mathbf{b}) \,,$$

where $\mathbf{u} \in \mathbb{R}^m$ is a Lagrange multiplier. By the Kuhn-Tucker necessary and sufficient conditions, a solution $\mathbf{x}^* = (x^*_j)$ of $A\mathbf{x} = \mathbf{b}$ is the optimal solution of (3.2) if and only if $\nabla_x\mathcal{L}(\mathbf{x}^*,\; \mathbf{u}) = \mathbf{0}$ for some $\mathbf{u}$. Let $\mathbf{x}^*$ be the solution of (3.2). Then

$$\nabla_x\mathcal{L}(\mathbf{x}^*,\; \mathbf{u}) = \mathbf{0} \iff N^{-p}(\mathbf{x}^* \circ |\mathbf{x}^*|^{p-2}) - A^T\mathbf{u} = \mathbf{0} \,. \tag{3.7}$$

Let the diagonal matrix $D = \mathrm{diag}\,(d_j)$ be defined by

$$d_j \;:=\; \begin{cases} \dfrac{n^p_j}{|x^*_j|^{p-2}} & , \text{ if } x^*_j \neq 0 \,, \\ 1 & , \text{ if } x^*_j = 0 \,. \end{cases}$$

Then (3.7) gives

$$\mathbf{x}^* = DA^T\mathbf{u} \,. \tag{3.8}$$

Substituting (3.8) into $A\mathbf{x} = \mathbf{b}$ gives

$$\mathbf{u} = (ADA^T)^{-1}\mathbf{b} \,.$$

Therefore

$$\mathbf{x}^* = DA^T(ADA^T)^{-1}\mathbf{b} \;\in \mathcal{X}_{\{2\}} \,. \tag{3.9}$$

Conversely, let $\mathbf{x}^*$ be any scaled minimum ℓ_2-norm solution, i.e.,

$$\mathbf{x}^* = DA^T\mathbf{u} \,, \quad \text{for some } D \in \mathcal{D}_n \,, \tag{3.10}$$

where $\mathbf{u} = (ADA^T)^{-1}\mathbf{b}$. Let the diagonal matrix $N = \mathrm{diag}\,(n_j)$ be defined by

$$n_j \;:=\; \begin{cases} \sqrt[p]{d_j|x^*_j|^{p-2}} & , \text{ if } x^*_j \neq 0 \,, \\ 1 & , \text{ if } x^*_j = 0 \,. \end{cases}$$

Then (3.10) gives

$$N^{-p}(\mathbf{x}^* \circ |\mathbf{x}^*|^{p-2}) - A^T\mathbf{u} = \mathbf{0} \,,$$

which, by (3.7), shows $\mathbf{x}^*$ to be the solution of (3.2). □

Theorem 3.2 Let $A \in \mathbb{R}_m^{m\times n}$. Then there is a solution $\mathbf{x}^*$ of

$$\min_{\mathbf{x}} \{ \ \|\mathbf{x}\|_1 : \ A\mathbf{x} = \mathbf{b} \ \} \tag{3.11}$$

which is a basic solution of $A\mathbf{x} = \mathbf{b}$, i.e., $\mathbf{x}^* = \widehat{A_{*J}^{-1}}\mathbf{b}$ for some $J \in \mathcal{J}$.

Proof. Let $\mathbf{y}$ be any solution of (3.11), and let $\mathbf{c} = \operatorname{sign}(\mathbf{y})$. Consider the linear programming problem

$$\text{(LP)} \qquad \begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b}, \\ & x_i \geq 0, \ \text{if } c_i = 1, \\ & x_i = 0, \ \text{if } c_i = 0, \\ & x_i \leq 0, \ \text{if } c_i = -1. \end{array}$$

Clearly $\mathbf{y}$ is an optimal solution of (LP), and any solution of (LP) is a solution of (3.11). By the theory of linear programming, there is a solution of (LP) which is a basic solution of $A\mathbf{x} = \mathbf{b}$. □

The following theorem is analogous to Lemma 1.1.

Theorem 3.3 Let $A \in \mathbb{R}_m^{m\times n}$ and let f be isotone. Then the problem

$$\min_{\mathbf{x}} \{ \ f(|\mathbf{x}|) \ : \ A\mathbf{x} = \mathbf{b} \ \} \ , \tag{3.12}$$

has a solution in $\mathcal{C}$. If f is strictly isotone then every solution of (3.12) lies in $\mathcal{C}$.

Proof. Let $\mathbf{x}^* = (x_i^*)$ be any solution of (3.12), and define three index sets for the signs of x_i^* ,

$$\pi = \{i \ : \ x_i^* > 0\} \ , \quad \zeta = \{i \ : \ x_i^* = 0\} \ , \quad \nu = \{i \ : \ x_i^* < 0\} \ .$$

Consider the polyhedral set

$$\mathcal{Y} = \{\mathbf{y} \ : \ A\mathbf{y} = \mathbf{b}, \ \mathbf{y}_\pi \geq 0, \ \mathbf{y}_\zeta = 0, \ \mathbf{y}_\nu \leq 0\} \ .$$

Since $\mathbf{x}^* \in \mathcal{Y}$, there exist extreme points $\mathbf{y}^{(1)}, \cdots, \ \mathbf{y}^{(r)}$ and extreme directions $\mathbf{d}^{(1)}, \cdots, \ \mathbf{d}^{(t)}$ of $\mathcal{Y}$ such that

$$\mathbf{x}^* = \sum_{i=1}^{r} \lambda_i \, \mathbf{y}^{(i)} + \sum_{j=1}^{t} \mu_j \, \mathbf{d}^{(j)} \ ,$$

where

$$\sum_{i=1}^{r} \lambda_i = 1, \quad \lambda_i \geq 0, \qquad \mu_j \geq 0 \ .$$

Moreover, the extreme points of $\mathcal{Y}$ are given by $\mathbf{y}^{(i)} = \widehat{A_{*J}^{-1}}\,\mathbf{b}$ for some $J \in \mathcal{J}$, and the extreme directions belong to the cone

$$\mathcal{D} = \{\mathbf{d} \ : \ A\mathbf{d} = 0, \ \mathbf{d}_\pi \geq 0, \ \mathbf{d}_\zeta = 0, \ \mathbf{d}_\mu \leq 0\} \ .$$

Let

$$\mathbf{x}^* = \mathbf{s} + \mathbf{d},$$

where

$$\mathbf{s} = \sum_{i=1}^{r} \lambda_i \, \mathbf{y}^{(i)}, \quad \mathbf{d} = \sum_{j=1}^{t} \mu_j \, \mathbf{d}^{(j)}.$$

Then

$$|\mathbf{x}^*| = |\mathbf{s}| + |\mathbf{d}|. \tag{3.13}$$

and

$$f(|\mathbf{x}^*|) \geq f(|\mathbf{s}|).$$

By the optimality of $\mathbf{x}^*$,

$$f(|\mathbf{x}^*|) = f(|\mathbf{s}|), \tag{3.14}$$

showing $\mathbf{s} \in \mathcal{C}$ is a solution of (3.12).
Next, suppose that f is strictly isotone. Then (3.14) implies $|\mathbf{x}^*| = |\mathbf{s}|$.

$$\therefore \quad \mathbf{d} = 0 \ , \quad \text{by (3.13)} \ . \qquad \therefore \quad \mathbf{x}^* = \mathbf{s} \in \mathcal{C} \ .$$ □

The following result, analogous to Theorems 2.4, 2.5, is stated without proof.

Theorem 3.4 $c\ell\, \mathcal{X}_{\{1\}} = c\ell\, \mathcal{X}_{\{2\}} = c\ell\, \mathcal{X}_{\{F\}}$, where

$$\mathcal{X}_{\{F\}} \ := \bigcup_{f \in \mathcal{F}_n} \left\{ \ \mathbf{x} \ : \ \mathbf{x} \in \arg\min_{\mathbf{x}} \{f(|\mathbf{x}|) \ : \ A\mathbf{x} = \mathbf{b}\} \ \right\} \ . \tag{3.15}$$

4 The general case

Throughout this section let $A \in \mathbb{R}_r^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. The basic solutions (0.7) are denoted

$$\mathbf{x}_{IJ} \ := \ \widehat{A_{IJ}^{-1}}\mathbf{b}_I \ , \quad (I, J) \in \mathcal{N} \ , \tag{4.1}$$

and their convex hull

$$\mathcal{C} \ := \ \operatorname{conv}\{\mathbf{x}_{IJ} \ : \ (I, J) \in \mathcal{N} \ \} \ . \tag{4.2}$$

Let $f_1, \ f_2$ be isotone functions. Consider the problem,

$$\min \{f_2(|\mathbf{x}|) \ : \ \mathbf{x} \in \arg \min_{\mathbf{x}} \ f_1(|A\mathbf{x} - \mathbf{b}|)\} \ . \tag{4.3}$$

For any full-rank factorization $A = CR$, clearly

$$\mathcal{I}(A) = \mathcal{I}(C) \ , \quad \mathcal{J}(A) = \mathcal{J}(R) \ , \tag{4.4}$$

and the above problem can be solved in stages:

$$\min_{\mathbf{y}} \ f_1(|C\,\mathbf{y} - \mathbf{b}|) \ , \tag{4.5}$$

$$\min_{\mathbf{x}} \{f_2(|\mathbf{x}|) \ : \ R\,\mathbf{x} = \mathbf{y}, \ \mathbf{y} \in \arg \min_{\mathbf{y}} \ f_1(|C\mathbf{y} - \mathbf{b}|)\} \ . \tag{4.6}$$

Combining Lemma 1.1 and Theorem 3.3 we have

Theorem 4.1 Let f_2 be isotone, and let f_1 be strictly isotone. Then there is a solution of (4.3) which is in $\mathcal{C}$. If in addition f_2 is strictly isotone, every solution of (4.3) lies in $\mathcal{C}$.

Proof. Let $A = CR$ be any full-rank factorization of A. Then $A_{IJ} = C_{I*}R_{*J}$. By Lemma 1.1 every solution $\mathbf{y}$ of (4.5) is a convex combination

$$\mathbf{y} = \sum_{I\in\mathcal{I}} \mu_I \, C_{I*}^{-1} \, \mathbf{b}_I \ . \qquad (4.7)$$

It follows from Theorem 3.3 that a solution of (4.6) is a convex combination

$$\begin{aligned} \mathbf{x} &= \sum_{J\in\mathcal{J}} \nu_J \, \widehat{R_{*J}^{-1}\mathbf{y}} \ , \\ &= \sum_{J\in\mathcal{J}} \nu_J \sum_{I\in\mathcal{I}} \mu_I \, \widehat{R_{*J}^{-1} C_{I*}^{-1}\mathbf{b}_I} \ , \quad \text{by (4.7)} \ , \\ &= \sum_{(I,J)\in\mathcal{N}} \lambda_{IJ} \, \mathbf{x}_{IJ} \ , \qquad (4.8) \end{aligned}$$

where

$$\lambda_{IJ} \ := \ \mu_I \, \nu_J \ , \ (I,J) \in \mathcal{N} \ , \qquad (4.9)$$

are also convex weights. The second part follows by applying the second part of Theorem 3.3. □

An immediate corollary of Theorem 4.1 is

Corollary 4.1 Let $1 \le p_1 < \infty$. Then the problem

$$\min \, \{ \|\mathbf{x}\|_{p_2} \ : \ \mathbf{x} \in \arg \min_{\mathbf{x}} \ \|A\mathbf{x} - \mathbf{b}\|_{p_1} \} \ , \qquad (4.10)$$

has a solution in $\mathcal{C}$. Moreover, if $1 \le p_2 < \infty$ then every solution of (4.10) lies in $\mathcal{C}$. □

The next example shows that Corollary 4.1 does not hold for $p_1 = \infty$.

Example 4.1 Let $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then the solution set of

$$\min_{\mathbf{x}} \ \|A\mathbf{x} - \mathbf{b}\|_\infty \qquad (4.11)$$

is $0 \le \mathbf{x} \le 2$. For all p, the minimum ℓ_p-norm best ℓ_∞-approximate solution is $\mathbf{x} = 0$, and does not belong to $\mathcal{C}$, which here is the singleton $\{1\}$.

References

[1] M.S. Bazaraa and C.M. Shetty, *Nonlinear Programming : Theory and Algorithms*, John Wiley & Sons, New York, 1979.

[2] A. Ben-Israel, "A volume associated with $m \times n$ matrices", *Lin. Algeb. Appl.* **167**(1992), 87-111.

[3] A. Ben-Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, Wiley-Interscience, 1974.

[4] A. Ben-Tal and M. Teboulle, "A geometric property of the least squares solution of linear equations", *Lin. Algeb. Appl.* **139**(1990), 165-170.

[5] L. Berg, "Three results in connection with inverse matrices", *Lin. Algeb. Appl.* **84**(1986), 63-77.

[6] M. Hanke and M. Neumann, "The geometry of the set of scaled projections", *Lin. Algeb. Appl.* (to appear)

[7] O.L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969.

[8] J. Miao and A. Ben-Israel, "On ℓ_p-approximate solutions of linear equations", *Lin. Algeb. Appl.* **199**(1994), 305–327