INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 10, Number 3, Pages 745–755

AN ALGORITHM FOR FINDING NONNEGATIVE MINIMAL NORM SOLUTIONS OF LINEAR SYSTEMS

SAÏD BAHI AND AARON ROSS

(Communicated by Edward Allen)

Abstract. A system of linear equations $Ax = b$, in n unknowns and m equations which has a nonnegative solution is considered. Among all its solutions, the one which has the least norm is sought when $\mathbb{R}^n$ is equipped with a strictly convex norm. We present a globally convergent, iterative algorithm for computing this solution. This algorithm takes into account the special structure of the problem. Each iteration cycle of the algorithm involves the solution of a similar quadratic problem with a modified objective function. Duality conditions for optimality are studied. Feasibility and global convergence of the algorithm are proved. As a special case we implemented and tested the algorithm for the ℓ^p-norm, where $1 < p < \infty$. Numerical results are included.

Key words. Linear equations, Least norms, Optimality, Duality conditions.

1. Introduction

We will be considering a system of real linear equations

$$Ax = b, \tag{1.1}$$

where A is an $m \times n$-real matrix, b in m-real vector and x an n-real vector. Under the assumption that the system (1.1) has a non-negative solution, we study the following problem: out of all non-negative solutions of (1.1) compute the solution that has the least norm when the norm considered on $\mathbb{R}^n$ is strictly convex. This naturally, includes the ℓ^p-norm, $1 < p < \infty$. The algorithm proposed in this work solves the minimal norm problem

$$\min \{\|x\| \mid Ax = b, \ x \in \mathbb{R}^n, \ x \geq 0\}. \tag{P}$$

We assume that $b \neq 0$, because otherwise the problem is trivial. All the steps of the algorithm for computing the solution of (P) will be shown to be feasible. Its global convergence will then be proved.

To solve the given problem, a dual problem, denoted (P'), will be associated with (P). An outline of the correspondence between (P) and (P') will be given. The main application of this work is the ℓ^p-norm case. Namely, find $x \in \mathbb{R}^n$ that minimizes the ℓ^p- problem

$$minimize\{ \|x\|_p \mid Ax = b, x \in \mathbb{R}^n, x \geq 0\}. \tag{1.2}$$

It should be noted here that the objective function in (1.2) need not be twice differentiable. The case $1 < p < 2$ has been more troublesome since methods requiring second derivatives will not be defined for certain non-zero points. While (1.2) is a smooth convex programming problem and thus susceptible to general programming procedures, it seems natural to take into account in our algorithm the special structure of the problem. For its convergence the proposed algorithm does not need extra differentiability of the norm.

Received by the editors July 5, 2012 and, in revised form, October 9, 2012.
2000 *Mathematics Subject Classification.* 90C25, 90C30.

For $p = 2$, problem (1.2) becomes a special case of what Lawson and Hanson referred to in [7] as the least distance programming (LDP) and for which they gave a finite algorithm. This algorithm (or any other similar purpose one) is used in this work as follows: at each iteration step of our algorithm, the LDP problem

$$Ax = b, x \geq 0, \quad \|x - a_k\|_2(min),$$

where (a_k) is a sequence defined by the algorithm, is solved using the LDP algorithm.

The main contribution of this paper is to propose a method of solution of probelm (P) that is not limited to a single norm such as the ℓ^2-norm. Different applications suggest different norms to use. Ideally, we seek a solution that optimize general norms. In many applications, a system of linear equations may have many solutions (e.g. when solving linear operator equations) and it may be needed, following a discretization, to select one solution under a given criteria. This criteria could be to find a solution that has the least norm or a solution that is the closest to a (target) point a in which case one needs to minimize $\|x - a\|$ among all solutions of a linear system. The classical ℓ^2-norm may not be always the best choice. For instence, in sparse solution construction and compressed sensing, similar ℓ^p minimization problems arise for $0 < p < 1$. Other applications arise when solving some variational problems.

2. Main notation and duality

Let the norm $\|\cdot\|$ on $\mathbb{R}^n, n \geq 1$, be arbitrary. The norm is said to be *smooth* if and only if through each point of unit norm there passes a unique hyperplane supporting the closed unit ball $B = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$. The norm is said to be *strictly convex* if and only if the unit sphere $S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ has no line segment on it.

To introduce the dual problem (P'), we define the dual norm $\|\cdot\|'$ on $\mathbb{R}^n$ by

$$\|y\|' = \max\{\langle x, y\rangle \mid \|x\| = 1,\ x \in \mathbb{R}^n\}.$$

For any vector $v \in \mathbb{R}^n, v \neq 0$, a $\|\cdot\|$-*dual vector*, v' is defined by

$$\|v'\| = 1,\ \langle v', v\rangle = \|v\|'. \tag{2.1}$$

Similarly, for the dual norm, a $\|\cdot\|'$-dual vector v^* is defined by

$$\|v^*\|' = 1,\ \langle v^*, v\rangle = \|v\|.$$

The map $y \mapsto y'$ (resp. $y \mapsto y^*$) is odd, continuous and positively homogeneous of degree zero on $\mathbb{R}^n\backslash\{0\}$, if the norm is strictly convex (resp. smooth). For $v \neq 0$, we have the relations $v^{'*} = v/\|v\|'$ (resp. $v^{*'} = v/\|v\|$) when the norm $\|\cdot\|$ is smooth (resp. strictly convex.)

When $\|\cdot\| = \|\cdot\|_p,\ 1 < p < \infty$, is the usual ℓ^p-norm, then $\|\cdot\|' = \|\cdot\|_q$, where $p + q = pq$. In terms of components, the dual vectors are given by

$$v_i' = (|v_i|/\|v\|_q)^{q-1} sgn\ (v_i), \quad v_i^* = (|v_i|/\|v\|_p)^{p-1} sgn\ (v_i), \quad i = 1, \ldots, n.$$

Let $K = \{x \in \mathbb{R}^n \mid x \geq 0,\ Ax = b\}$. Given problem (P), we associate a dual problem ([1], [8])

$$(P') \qquad \max\{\langle b, y\rangle \mid\ \xi \in \mathbb{R}^n, \xi \geq 0,\ y \in \mathbb{R}^m,\ \|\xi + A^T y\|' \leq 1\},$$

where A^T is the transpose of the matrix A. The relation between (P) and (P') is studied in the next two results.