LEAST SQUARES PROBLEM
FOR LINEAR INEQUALITIES WITH BOUNDS ON THE VARIABLES

ELENA POPESCU*

Abstract. In [2] S.P. Han proposed a method for finding a least-squares solution for systems of linear inequalities. Additional information about the real world problem may take the form of additional inequalities. A common case is when the variables are restricted to lie in certain prescribed intervals. A method for this very important case is the subject of this paper.

MSC 2000. 90C25.
Keywords. Convex programming, box constraints.

1. INTRODUCTION

Consider the system of linear inequalities

$$(1) \qquad Ax \leq b,$$

where $A \in \mathcal{M}_{m \times n}$, $b \in \mathbb{R}^m$, $x \in \mathbb{R}^n$. When the system is inconsistent, we are interested in vectors that satisfy the system (1) in a least-squares (LS) sense, that is, vectors $x \in \mathbb{R}^n$ solving

$$(2) \qquad \min \tfrac{1}{2}\|(Ax - b)_+\|_2^2,$$

where $\|\cdot\|$ stands for the Euclidean norm in $\mathbb{R}^m$ and $(Ax - b)_+$ is the m-vector whose i-th component is $\max\{(Ax - b)_i, 0\}$. The linear least squares problem of equalities has always enjoyed a lot of attention from researchers (see [1]).

We are concerned with LS problem (2) of inequalities when the components of x are within meaningful intervals:

$$(3) \qquad l_i \leq x_i \leq u_i, \quad i = 1, ..., n.$$

A vector which verifies the constraints (3) will be called a *feasible point*.

In paper [3] we have worked out an algorithm for solving convex programs subject to box constraints. In this paper we have applied the algorithm to solve the problem (2) with (3).

The method proposed proceeds by solving a finite number of smaller unconstrained subproblems, such a subproblem having the form of least squares

---

*University "Ovidius", B-dul Mamaia nr. 124, Ro-900527 Constanţa, Romania, e-mail: epopescu@univ-ovidius.ro.

minimization over a linear subspace. Is an adaptive method for determining the constraints that are active at optimal solution $\widehat{x}$, i.e. the components of $\widehat{x}$ that are exactly at one of their bounds. The method considered develops on two levels. At each iteration of a higher level (it will be called *major iteration*) it is decided which variables are fixed at one of its bounds and which one are free, that is, strictly between its bounds.

At lower level, a least squares subproblem is solved at a time, only in the subspace of the free variables while keeping the fixed variables unchanged. The iterations of this subproblem will be called *minor iterations*. The starting point for the minor iterations is a feasible point given by the current major iteration.

Consider the following problem:

(4)
$$\begin{cases} \min f(x) \\ l_i \leq x_i \leq u_i, \quad i = 1, ..., n, \end{cases}$$

where f is the convex differentiable function

(5)
$$f(x) = \tfrac{1}{2}(Ax - b)_+^{\mathrm{T}}(Ax - b)_+.$$

Its gradient is

(6)
$$\nabla f(x) = A^{\mathrm{T}}(Ax - b)_+.$$

For any feasible point x, we use $N_l(x)$ and $N_u(x)$ to denote the sets of indices for which the corresponding components of the point x are fixed at one of its bounds, that is:

$$N_l(x) = \{i : x_i = l_i\}$$

$$N_u(x) = \{i : x_i = u_i\}$$

and $N(x) = N_l(x) \cup N_u(x)$.

If $\widehat{x}$ is an optimal point for problem (4) then it is also optimal for "restraint" problem:

$$\begin{cases} \min f(x) \\ l_i \leq x_i \leq u_i, \quad i \in N(\widehat{x}). \end{cases}$$

If we can determine the sets $N_l(\widehat{x})$ and $N_u(\widehat{x})$, we can solve the "restraint" problem with constraints taken as equalities:

$$\begin{cases} \min f(x) \\ x_i = l_i, \quad i \in N_l(\widehat{x}) \\ x_i = u_i, \quad i \in N_u(\widehat{x}). \end{cases}$$

Consider the subproblem

(7)
$$\begin{cases} \min f(x) \\ x_i = l_i, \quad i \in N_l^p \\ x_i = u_i, \quad i \in N_u^p. \end{cases}$$

Assume that F^p is the complement of $N^p = N_l^p \cup N_u^p$, that is, the set of indices of free variables at major iteration p. We also assume without loss of generality, that

$$F^p = \{1, ..., q\}.$$

Then we have the following partitions:

$$x = \begin{pmatrix} z \\ \bar{z} \end{pmatrix}, \quad A = (C\ D),$$

where the subvector z contains the first q components of x (the free part), while $\bar{z}$ contains the last $n - q$ components (the fixed part). The matrix $C \in \mathcal{M}_{m \times q}$ contains the first columns of A.

The subproblem (7) then reduces to finding a vector $z^{(p)} \in \mathbb{R}^q$ solving

$$(8) \qquad\qquad \min_{z \in \mathbb{R}^q} \tfrac{1}{2} \| (Cz - d)_+ \|_2^2$$

where

$$d = b - D\bar{z}.$$

From practical point of view, (8) is an unconstrained problem with q free variables.

Then, returning to the n-dimensional space, the optimal point for (7) is

$$y^p = \begin{pmatrix} z^{(p)} \\ \bar{z} \end{pmatrix}.$$

2. SOLVING THE SUBPROBLEM IN THE FREE VARIABLE

For solving the least squares subproblem (8), we have chosen to use the S.-P. Han algorithm [2] with modification that the iterations of the algorithm (minor iterations) are stopped as soon as one or more of the free variables violate their bounds. The starting point for minor iterations is given by the current major iteration.

The Han's method has a finite convergence property.

The function

$$g(z) = \tfrac{1}{2} \| (Cz - d)_+ \|_2^2$$

is differentiable and convex and its gradient is:

$$\nabla g(z) = C^{\mathrm{T}} (Cz - d)_+.$$

The necessary and sufficient condition for an q-vector $\widehat{z}$ to be a least squares solution of the system

$$(9) \qquad\qquad\qquad Cz \leq d$$

is

$$(10) \qquad\qquad\qquad \nabla g(\widehat{z}) = 0.$$

Condition (10) may be called the *normal equation* of the system (9) because it is analogous to the normal equation for the equality case.

In [2] it is pointed that a least squares solution of (9) exists. The system (9) may be also written as

$$c_i^{\mathrm{T}} z \leq d_i, \quad i \in M = \{1, 2, ..., m\}$$

where c_i^{T} forms the i-th row of the matrix C and d_i the i-th component of d. For $z \in \mathbb{R}^q$ we use $I(z)$ to denote the set of indices $i \in M$ of active and violated inequalities at z. That is,

$$I(z) = \{i : c_i^{\mathrm{T}} z \geq d_i\}.$$

For $I = I(z)$ we denote C_I the submatrix of C that consists of rows whose indices are in I, while the subvector d_I is similarly defined.

ALGORITHM 1 (least squares solution of linear inequalities).
1. Initialization: $z^{(0)} \in \mathbb{R}^q$, ($z^{(0)}$ *contains the free part of current major iteration x^p*).
2. Typical step: *Generate the direction*

$$\bar{w} = C_I^+ (d_I - C_I z^{(j)}),$$

where $I = I(z^{(j)})$ and C_I^+ is the pseudo-inverse of C_I. Determine the stepsize $\bar{\lambda}$-the smallest minimizer of function:

$$\theta(\lambda) = g(z^{(j)} + \lambda \bar{w}).$$

Set

$$z^{(j+1)} = z^{(j)} + \bar{\lambda} \bar{w}.$$

In [2] it is pointed that the direction $\bar{w}$ is a descendent direction for the function g at $z^{(j)}$, more precisely:

$$\nabla g(z^{(j)}) = -C_I^{\mathrm{T}} C_I \bar{w}.$$

If $\bar{w} = 0$, then from (10) $z^{(j)}$ is a least squares solution of (9). Else the process is repeated until

$$\|C^{\mathrm{T}}(Cz^{(j)} - d)_+\| \leq \varepsilon$$

or one or more of the components of $z^{(j)}$ violate their bounds.

The pseudo-inverse can be constructed by a singular value decomposition. In any case, to compute $\bar{w}$ it is not necessary to compute the pseudo-inverse explicitly.

3. THE METHOD

Let y^p be optimal for problem (7). It follows that there are Lagrange multipliers λ_i, $i \in N^p = N_l^p \cup N_u^p$ such that:

$$(11) \qquad \nabla f(y^p) + \sum_{i \in N_l^p} \lambda_i e^i + \sum_{i \in N_u^p} \lambda_i e^i = 0,$$

i.e. $\nabla f(y^p)$ is expressed as unique linear combination of unit vectors $e^i = (0, ..., 1^i, ..., 0)$.

If conditions

$$(12) \qquad l_i \leq y_i^p \leq u_i, \; i \notin N^p,$$

$$(13) \qquad \begin{cases} \lambda_i \leq 0, & i \in N_l^p \\ \lambda_i \geq 0, & i \in N_u^p \end{cases}$$

are satisfied, then y^p is optimal solution for problem (4) (see [3]).

If

$$L = \min_{i \in N^p} \alpha_i \lambda_i$$

with

$$\alpha_i = \begin{cases} -1, & \text{if } i \in N_l^p \\ 1, & \text{if } i \in N_u^p \end{cases}$$

then condition (13) may be written as $L \geq 0$.

From (6) and (11), Lagrange multipliers are calculated as it follows:

$$\lambda_i = -\{A^{\mathrm{T}}(Ay^p - b)_+\}_i, \quad i \in N^p.$$

The method is briefly the following: suppose the starting feasible point x^1 is given; then we solve the subproblem

$$(14) \qquad \begin{cases} \min f(x) \\ x_i = l_i, \; i \in N_l(x^1) \\ x_i = u_i, \; i \in N_u(x^1) \end{cases}$$

and let y^1 be optimal point for problem (14). We distinguish between two possibilities.

The first case: y^1 is feasible for problem (2)–(3) (i.e. $l_i \leq y_i^1 \leq u_i, \; i \notin N(x^1)$); then:

$$x^2 = y^1.$$

If y^1 is also optimal for problem (2)–(3), (i.e. $L \geq 0$), the procedure stops.

If $L < 0$ then a subspace of a large dimension in considered, by releasing a constraint (according to Proposition 2 and 3, from [3], a fixed variable leaves a bound and becomes free) and a smaller value of f will be obtained.

The second case: y^1 is not feasible for problem (2)–(3) (there exists $i \notin N(x^1)$ such that $y_i^1 < l_i$ or $y_i^1 > u_i$), then let x^2 be the feasible point closest to y^1 on the line segment between x^1 and y^1:

$$x^2 = x^1 + \tau(y^1 - x^1),$$

where the scalar $\tau \in (0,1)$ in determined such that

$$l_i \leq x_i^2 \leq u_i, \; i = 1, ..., n$$

and x^2 to be closest y^1. At least one additional component $r \notin N(x^1)$ of x^2 hits one of their bounds. Then we solve the subproblem:

$$\begin{cases} \min f(x) \\ x_i = l_i, & i \in N_l(x^2) \\ x_i = u_i, & i \in N_u(x^2). \end{cases}$$

ALGORITHM 2 (suboptimization algorithm for LS problem with interval constraints).

1. Initialization: $x^1 \in \mathbb{R}^n$ a feasible point, $N_l^1 = N_l(x^1)$, $N_u^1 = N_u(x^1)$ and $N^1 = N_l^1 \cup N_u^1$ the sets of indices of the components of x^1 that are exactly at one of their bounds.

2. Typical Step: Solve the subproblem on the free variables keeping the fixed variables

$$(15) \qquad \begin{cases} x_i = l_i, & i \in N_l^p \\ x_i = u_i, & i \in N_u^p \end{cases}$$

unchanged, that is, determine $z^{(p)}$ optimal for

$$\min_{z \in \mathbb{R}^q} \tfrac{1}{2}\|(Cz - d)_+\|_2^2.$$

Determine

$$y^p = \begin{pmatrix} z^{(p)} \\ \bar{z} \end{pmatrix},$$

where $\bar{z}$ contains the fixed part (15).

Case 1: y^p is a feasible point

Determine

$$L = \min_{i \in N^p} \alpha_i \lambda_i$$

with

$$\alpha_i = \begin{cases} -1, & \text{if } i \in N_l^p \\ 1, & \text{if } i \in N_u^p \end{cases}$$

where λ_i are Lagrange multipliers associated with the bound constraints:

$$\lambda_i = -\{A^{\mathrm{T}}(Ay^p - b)_+\}_i, \ i \in N^p.$$

If $L \geq 0$, STOP; y^p is optimal for problem (2)–(3).

If $L < 0$, let $i_0 \in N^p$ such that $L = \alpha_{i_0}\lambda_{i_0}$.

Define

$$N_l^{p+1} = \begin{cases} N_l^p \setminus \{i_0\}, & \text{if } i_0 \in N_l^p \\ N_l^p, & \text{if } i_0 \notin N_l^p \end{cases}$$

$$N_u^{p+1} = \begin{cases} N_u^p \setminus \{i_0\}, & \text{if } i_0 \in N_u^p \\ N_u^p, & \text{if } i_0 \notin N_u^p \end{cases}$$

$$N^{p+1} = N^p \setminus \{i_0\}$$

$$x^{p+1} = y^p.$$

Go to 2 with $p + 1$ replacing p.

Case 2: y^p is not a feasible point.

Determine x^{p+1} the feasible point closest to y^p on the line between x^p and y^p. Define

$$N_l^{p+1} = N_l(x^{p+1})$$

$$N_u^{p+1} = N_u(x^{p+1})$$
$$N^{p+1} = N_l^{p+1} \cup N_u^{p+1}.$$

Go to 2 *with* $p + 1$ *replacing* p.

The algorithm is a finite process because solves a finite number of subproblems and each subproblem is solved in a finite number of iterations.

REFERENCES

[1] Bierlaire M., Toint Ph. L., and Tuyttens D., *On iterative algorithms for linear least squares problems with bound constraints*, Linear Algebra Appl., **143**, pp. 111–143, 1991.

[2] Han S. P., *Least-squares solution of linear inequalities*, Technical Report 2141, Mathematics Research Center, University of Wisconsin-Madison, 1980.

[3] Stancu-Minasian I. M. and Popescu E., *Algorithm for solving convex programs subject to box constraints*, Analele Universităţii Bucureşti, Matematica, **XLIX**, no. 1, pp. 99–108, 2000.