MATHEMATICA — REVUE D'ANALYSE NUMÉRIQUE
ET DE THÉORIE DE L'APPROXIMATION
L'ANALYSE NUMÉRIQUE ET LA THÉORIE DE L'APPROXIMATION
Tome 9, N° 2, 1980, pp. 209—219

ON A GEOMETRIC PROGRAMMING PROBLEM WITH LINEAR CONSTRAINTS

by
I. MARUȘCIAC
(Cluj-Napoca)

1. Introduction

In this paper a special class of geometric programming problems is considered. In order to solve such a problem, i.e. to minimize a general posinomial subject to arbitrary linear constraints, a variant of conjugate gradients method is proposed. First a new form of optimal criteria for the solution of such a problem is established. Then a method of projection gradients are described to slove the problem. To illustrate the algorithm a small example are also presented.

2. Preliminary results

In this section some results regarding the optimal solution of the geometric programs with linear constraints are established. The problem considered here is the following: given $p : \overset{0}{\mathbf{R}}{}^n_+ \to \mathbf{R}$,

$$p(\mathbf{x}) = \sum_{i \in I} c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}},$$

where $a_{ij} \in \mathbf{R}$, $i \in I$, $j = 1, 2, \ldots, n$, $c_i \in \overset{0}{\mathbf{R}}_+$, $i \in I$, called *posinomial*, and

$$\Omega = \{\mathbf{x} \in R^n : \mathbf{Bx} \leqslant \mathbf{b},\ \mathbf{x} > \mathbf{0}\},$$

where $\mathbf{B} = (b_{ij}) \in \mathbf{M}_{m\times n}(\mathbf{R})$, $\mathbf{b} \in \mathbf{M}_{m\times 1}(\mathbf{R})$, the problem we will dealing with is:

$$(1) \qquad \inf\,\{p(\mathbf{x}) : \mathbf{x} \in \Omega\}$$

or explicitely

$$(1') \qquad \inf\{\sum_{i\in I} c_i x_1^{a_{i1}} x_2^{a_{i2}} \ \ldots\ x_n^{a_{in}} : \ \mathbf{Bx} \leqslant \mathbf{b},\ \mathbf{x} > 0\}.$$

Consider $\mathfrak{g} : \mathbf{R}^n \to \mathbf{R}^n_+$,

$$\mathfrak{g}(\mathbf{z}) = (e^{z_1}\ e^{z_2}\ \ldots\ e^{z_n})^T = \mathbf{e}^{\mathbf{z}}.$$

If we take $\mathbf{x} = \mathbf{e}^{\mathbf{z}}$, then

$$q(\mathbf{z}) = (p \circ \mathfrak{g})(\mathbf{z}) = \sum_{i\in I} c_i e^{\mathbf{a}^{i\cdot}\mathbf{z}},$$

where

$$\mathbf{a}^{i\cdot} = (a_{i1} a_{i2}\ \ldots\ a_{in}) \in \mathbf{M}_{1\times n}(\mathbf{R}),$$

is a convex function on $\mathbf{R}^n$, and problem (1) is transformed into the following equivalent one:

$$(2) \qquad \inf\{\sum_{i\in I} c_i e^{\mathbf{a}^{i\cdot}\mathbf{z}} : \mathbf{B}e^{\mathbf{z}} \leqslant \mathbf{b}\}.$$

LEMMA 1. *$\mathbf{x}^0 \in \mathbf{R}^n$ is optimal solution to (1) iff $\mathbf{z}^0 = \ln \mathbf{x}^0 = (\ln x_1^0, \ln x_2^0, \ldots, \ln x_n^0) \in \mathbf{R}^n$ is optimal solution to (2).*

Proof. ($\Rightarrow$) Consider $\mathbf{x}^0 \in \Omega$ optimal solution to (1). Then $\mathbf{z}^0 = \ln \mathbf{x}^0$ is optimal to (2). Indeed, since $\mathbf{x}^0 \in \Omega$ we have

$$\mathbf{Ax}^0 \leqslant \mathbf{b},\ \mathbf{x}^0 > \mathbf{0},$$

or

$$\mathbf{A}e^{\mathbf{z}^0} \leqslant \mathbf{b},$$

and, therefore, $\mathbf{z}^0$ is a feasible solution to (2). Assume the contrary, that $\mathbf{z}^0$ is not optimal solution to (2). Then there is $\mathbf{z}' \in \mathbf{R}^n$, for which $\mathbf{A}e^{\mathbf{z}'} \leqslant \mathbf{b}$ and such that

$$q(\mathbf{z}') < q(\mathbf{z}^0),$$

that means

$$p(\mathfrak{g}(\mathbf{z}')) < p(\mathfrak{g}(\mathbf{z}^0)),$$

or

$$p(\mathbf{x}') < p(\mathbf{x}^0),\ \ \mathbf{x}' = \mathbf{e}^{\mathbf{z}'},$$

a contradiction.

($\Leftarrow$) Is similar.

In the case when $b_{ij} \in \mathbf{R}_+$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, problem (2) is obviously a convex programming problem.

Denote by

$$Z = \{\mathbf{z} \in \mathbf{R}^n : \mathbf{B}e^{\mathbf{z}} \leqslant \mathbf{b}\}$$

the feasible set of the problem (2).

3. Necessary and sufficient conditions for optimal solution

If $\mathbf{z}^0 \in Z$, let $J_0 = \{i : \mathbf{b}^{i\cdot}\mathbf{e}^{\mathbf{z}^0} = b_i\}$ and $J_- = \{i : \mathbf{b}^{i}\mathbf{e}^{\mathbf{z}^0} < \mathbf{b}_i\}$ ($J_- \cup J_0 = \{1, 2, \ldots, m\}$).

Assume that vectors $\mathbf{b}^{i\cdot}$, $i \in J_0$ are linearly independent. Then if $\mathbf{B}_{J_0}$ is the matrix formed by the vectors $\mathbf{b}^{i\cdot}$, $i \in J_0$, then (see [2], p. 147) matrices:

$$\mathbf{Q} = \mathbf{B}_{J_0}^T(\mathbf{B}_{J_0}\mathbf{B}_{J_0}^T)^{-1}\mathbf{B}_{J_0} \in \mathbf{M}_{n\times n}(\mathbf{R})$$

$$\mathbf{P} = E - \mathbf{Q},$$

are the operators of orthogonal projection onto the subspace $\mathbf{D} \subset \mathbf{R}$ spanned by the vectors $\mathbf{b}^{i\cdot}$, $i \in J_0$ and $\mathbf{D}^{\perp}$ — the orthogonal subspace of $\mathbf{D}$, respectively.

It is known that

$$\mathbf{QQ} = \mathbf{Q}; \quad \mathbf{PP} = \mathbf{P}$$

$$\mathbf{Q}^T = \mathbf{Q}; \quad \mathbf{P}^T = \mathbf{P}$$

$$\mathbf{PQ} = \mathbf{0}.$$

LEMMA 2. *Assume that $\mathbf{B} \in \mathbf{M}_{m\times n}(\mathbf{R}_+)$, $\mathbf{b} \in \mathbf{M}_{m\times 1}(\mathbf{R})$ and that (2) is superconsistent. Then $\mathbf{z}^0 \in Z$ is optimal solution to the problem (2) iff*

$$\text{(i)} \qquad \sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{P}\mathbf{a}^{i\cdot T}(\mathbf{z}^0) = \mathbf{0}$$

$$\text{(ii)} \qquad \sum_{i\in I} v_i(\mathbf{z}^0)(\mathbf{B}_J\mathbf{B}_{J_0}^T)^{-1}\mathbf{B}_{J_0}\mathbf{a}^{i\cdot T}(\mathbf{z}^0) \leqslant \mathbf{0},$$

where

$$v_i(\mathbf{z}) = c_i e^{\mathbf{a}^{i\cdot}(\mathbf{z})},$$

$$\mathbf{a}^{i\cdot}(\mathbf{z}) = \left(\frac{a_{i1}}{e^{z_1}}, \frac{a_{i2}}{e^{z_2}}, \ldots, \frac{a_{in}}{e^{z_n}}\right)$$

Proof. Since in this case problem (2) is convex and superconsistent with continuously differentiable objective and constraint functions, from Kuhn—Tucker's theorem it follows that $\mathbf{z}^0 \in Z$ is optimal solution to (2) iff there is $\mathbf{u}^0 \in \mathbf{R}_+^m$ such that

1^0 $u_i^0(\mathbf{b}^{i\cdot}\mathbf{e}^{\mathbf{z}^0} - b_i) = 0, \; i = 1, 2, \ldots, m;$

2^0 $\nabla q(\mathbf{z}^0) + \sum_{i=1}^{m} u_i^0 \nabla(\mathbf{b}^{i\cdot}\mathbf{e}^{\mathbf{z}^0} - b_i) = \mathbf{0}.$

Since $u_i^0 = 0$, $i \in J_-$, from 2^0 we have

$$\sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{a}^{i\cdot}(\mathbf{z}^0) + \sum_{i\in J_0} u_i^0 \mathbf{b}^{i\cdot} = \mathbf{0}, \tag{3}$$

From (3) it is seen that

$$-\Big(\sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{a}^{i\cdot}(\mathbf{z}^0)\Big) \in \mathbf{D}^{\perp}$$

and, therefore,

$$\mathbf{P}\Big(\sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{a}^{i\cdot T}(\mathbf{z}^0)\Big) = \sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{P}\mathbf{a}^{i\cdot T}(\mathbf{z}^0) = \mathbf{0},$$

i.e. (i) holds.

In order to prove (ii) we observe that (3) can be written under the from

$$\sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{a}^{i\cdot}(\mathbf{z}^0) + \mathbf{u}^{0T}\mathbf{B}_{J_0} = \mathbf{0}$$

or, by transposing,

$$\sum_{i\in I} v_i(\mathbf{z}^0)\mathbf{a}^{i\cdot T}(\mathbf{z}^0) + \mathbf{B}_{J_0}\mathbf{u}^0 = 0. \tag{4}$$

Multiplying (4) by $(\mathbf{B}_{J_0}\mathbf{B}_{J_0}^T)^{-1}\mathbf{B}_{J_0}^T$ we get

$$-\mathbf{u}^0 = \sum_{i\in I} v_i(\mathbf{z}^0)(\mathbf{B}_{J_0}\mathbf{B}_{J_0}^T)^{-1}\mathbf{B}_{J_0}\mathbf{a}^{i\cdot T}(\mathbf{z}^0) \leqslant \mathbf{0},$$

i.e. (ii).

THEOREM 1. *Let* (1) *be a superconsistent geometric program with* $b_{ij} \in \mathbf{R}_+$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$. *Then* $\mathbf{x}^0 \in \Omega$ *is optimal solution to* (1) *if and only if*

$1°$ $$\sum_{i\in I} u_i(\mathbf{x}^0)\mathbf{P}\mathbf{a}^{i\cdot T}(\mathbf{x}^0) = \mathbf{0},$$

$2°$ $$\sum_{i\in I} u_i(\mathbf{x}^0)(\mathbf{B}_{J_0}\mathbf{B}_{J_0}^T)^{-1}\mathbf{B}_{J_0}\mathbf{a}^{i\cdot T}(\mathbf{x}^0) \leqslant \mathbf{0},$$

where

$$u_i(\mathbf{x}) = c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}}, \; i \in I,$$

$$\mathbf{a}^{i\cdot}(\mathbf{x}) = \left(\frac{a_{i1}}{x_1^0}, \frac{a_{i2}}{x_2^0}, \ldots, \frac{a_{in}}{x_n^0}\right)$$

Proof. Lemma 1 shows that $\mathbf{x}^0$ is optimal solution to (1) if and only if $\mathbf{z}^0 = \ln \mathbf{x}^0$ is optimal solution to (2). As program (1) is superconsistent so is (2). Since $b_{ij} \in \mathbf{R}_+$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$, program (2) is convex, and from Lemma 2, $\mathbf{z}^0$ is optimal for (2) if and only if (i)—(ii) hold.

Because $\mathbf{z}^0 = \ln \mathbf{x}^0$, we have

$$v_i(\mathbf{z}^0) = c_i e^{\sum_{j=1}^{n} a_{ij} z_j^0} = c_i e^{\sum_{j=1}^{n} a_{ij} \ln x_j^0} = c_i \prod_{j=1}^{n} (x_j^0)^{a_{ij}} = u_i(\mathbf{x}^0);$$

$$\mathbf{a}^{i\cdot}(z^0) = \left(\frac{a_{i1}}{e^{z_1^0}}, \frac{a_{i2}}{e^{z_2^0}}, \ldots, \frac{a_{in}}{e^{z_n^0}}\right) = \left(\frac{a_{i1}}{x_1^0}, \frac{a_{i2}}{x_2^0}, \ldots, \frac{a_{in}}{x_n^0}\right) = \mathbf{a}^{i\cdot}(\mathbf{x}^0).$$

Therefore (i)—(ii) are equivalent to 1^0—2^0.

4. Minimization of a posinomial on a subspace

Now suppose that we have to minimize the posinomial

$$p(\mathbf{x}) = \sum_{i\in I} u_i(\mathbf{x}) = \sum_{i\in I} c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}} \tag{5}$$

subject to the liniar constraints

$$\mathbf{b}^{i\cdot}\mathbf{x} = b_i, \; i \in J. \tag{6}$$

Assume that vectors $\mathbf{a}^{i\cdot}$, $i \in J$, are linearly independent. Let $\mathbf{x}^0 \in \mathbf{R}_+^n$ be a point which satisfies (6), i.e.

$$\mathbf{B}_J\mathbf{x}^0 = \mathbf{b}_J,$$

where $\mathbf{b}_J$ is a vector whose components are b_i, $i \in J$.

Now we introdus a new variable $\mathbf{y} \in \mathbf{R}^n$ defined as follows:

$$\mathbf{x} = \mathbf{x}^0 + \mathbf{P}\mathbf{y}, \quad \mathbf{P} = \mathbf{E} - \mathbf{B}_J^T(\mathbf{B}_J\mathbf{B}_J^T)^{-1}\mathbf{B}_J,$$

and consider the function $\varphi : \mathbf{R}^n \to \mathbf{R}^n$,

$$\varphi(\mathbf{y}) = p(\mathbf{x}^0 + \mathbf{P}\mathbf{y}).$$

LEMMA 3. *If* $\mathbf{x}^0 \in X = \{\mathbf{x} \in \mathbf{R}_+^n : \mathbf{B}_J\mathbf{x} = \mathbf{b}_J\}$, *then*

$$\forall \mathbf{y} \in \mathbf{R}^n \Rightarrow \mathbf{x} = \mathbf{x}^0 + \mathbf{P}\mathbf{y} \in X.$$

Proof. First we observe that

$$\mathbf{B}_J\mathbf{P} = \mathbf{B}_J(E - \mathbf{B}_J^T(\mathbf{B}_J\mathbf{B}_J^T)^{-1}\mathbf{B}_J) = \mathbf{B}_J - (\mathbf{B}_J\mathbf{B}_J^T)(\mathbf{B}_J\mathbf{B}_J^T)^{-1}\mathbf{B}_J = \mathbf{0},$$

thus

$$\mathbf{B}_J\mathbf{x} = \mathbf{B}_J(\mathbf{x}^0 + \mathbf{P}\mathbf{y}) = \mathbf{B}_J\mathbf{x}^0 + \mathbf{B}_J\mathbf{P}\mathbf{y} = \mathbf{B}_J\mathbf{x}^0 = \mathbf{b}_J.$$

According to the rules of differentiation of a composite function and in view of the symmetry of operator $\mathbf{P}$, we have for each $\mathbf{y} \in \overset{0}{\mathbf{R}}{}_+^n$ such that $\mathbf{x}^0 + \mathbf{P}\mathbf{y} \in \overset{0}{\mathbf{R}}{}_+^n$,

(7) $$\nabla^T\varphi(\mathbf{y}) = \mathbf{P}\nabla^T p(\mathbf{x}),$$

THEOREM 2. *Assume that* $b_{ij} \in \mathbf{R}_+$, $i \in J$, $j = 1, 2, \ldots, n$. *If* $\mathbf{y}^* \in \mathbf{R}$ *is a point of global minimum of the function* φ *and the corresponding point*

$$\mathbf{x}^* = \mathbf{x}^0 + \mathbf{P}\mathbf{y}^* \in \overset{0}{\mathbf{R}}{}_+^n,$$

then $\mathbf{x}^*$ *is the minimum point of* p *on* X.

Proof. Let $y^* \in \mathbf{R}$ be a point of global minimum of φ. Since in this case φ is differentiable at $\mathbf{y}^*$ (as a composite of two differentiable functions), it follows that

$$\nabla\varphi(\mathbf{y}^*) = \mathbf{0},$$

therefore

(8) $$\mathbf{P}\nabla^T p(\mathbf{x}^*) = \mathbf{0}.$$

Taking

$$\mathbf{u} = -(\mathbf{B}_J\mathbf{B}_J^T)^{-1}\mathbf{B}_J\nabla^T p(\mathbf{x}^*),$$

from (8) we derive

$$(\mathbf{E} - \mathbf{B}_J^T(\mathbf{B}_J\mathbf{B}_J^T)^{-1}\mathbf{B}_J)\nabla^T p(\mathbf{x}^*) = \nabla^T p(\mathbf{x}^*) + \mathbf{B}_J^T\mathbf{u} = \mathbf{0},$$

which represents the necessary condition for $\mathbf{x}^*$ to be minimum point to p on X.

But condition

(9) $$\nabla^T p(\mathbf{x}^*) + \mathbf{B}_J^T\mathbf{u} = \mathbf{0}$$

is equivalent to

(10) $$\frac{\partial p(x^*)}{\partial xj} + \sum_{i \in J} u_i b_{ij} = 0, \; j = 1, 2, \ldots, n.$$

From

$$p(\mathbf{x}) = q(\ln \mathbf{x}),$$

we have

(11) $$\text{(11)} \; \frac{\partial p(x^*)}{\partial x_j} = \frac{\partial q(z^*)}{\partial z_j} \cdot \frac{1}{x_j^*}.$$

Replacing (11) in (10) we get

$$\frac{\partial q(z^*)}{\partial z_j} + \sum_{i \in J} u_i b_{ij} e^{z_j^*} = 0, \; j = 1, 2, \ldots, n,$$

which shows that $\mathbf{z}^* = \ln \mathbf{x}^*$ is a minimum point of q on

$$Z = \{\mathbf{z} \in \mathbf{R}^n : \mathbf{B}_J e^{\mathbf{z}} = \mathbf{b}_J\},$$

since q and constraints

$$\sum_{j=1}^{n} b_{ij} e^{z_j} - b_i = 0, \quad i \in J$$

are convex.

Now from Theorem 1 it folliws that $\mathbf{x}^*$ is also a minimum point of p on X.

Theorem 2 shows that the problem

$$\min\{p(\mathbf{x}) : \mathbf{B}_J\,\mathbf{x} = \mathbf{b}_J, \; \mathbf{x} > \mathbf{0}\}$$

can be reduced to the minimization of φ without constraints.

To minimize $\varphi(\mathbf{y})$ we apply the method of conjugate gradients:

$$\mathbf{y}^0 = \mathbf{0}$$

$$\mathbf{d}^1 = -\nabla^T\varphi(\mathbf{0})$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \alpha_{k+1}\mathbf{d}^{k+1}, \quad k \geqslant 0$$

(12) $$\mathbf{d}^{k+1} = -\nabla^T\varphi(\mathbf{y}^k) + \frac{\|\nabla\varphi^k)\|^2}{\|\nabla\varphi(\mathbf{y}^{k-1})\|^2}\mathbf{d}^k$$

(13) $$\alpha_{k+1} = -\frac{\nabla(\mathbf{y}^k)\mathbf{d}^{k+1}}{\mathbf{d}^{k+1^T}\nabla^2\varphi(\mathbf{y}^k)\mathbf{d}^{k+1}}$$

LEMMA 4. *For each* $k \in N$,

(14) $$\mathbf{P}\mathbf{d}^k = \mathbf{d}^k$$

Proof. Can be done by induction, using (12), (13), $\mathbf{PP} = \mathbf{P}$ and the relationship between $\nabla\varphi(\mathbf{y})$ and $\nabla p(\mathbf{x})$.

THEOREM 3. *The problem of minimization of the posinomial* p *with constraints* (6) *is solved by the following algorithm : given an initial point* x^0 *which satisfies* (6),

$$\mathbf{d}^1 = -\sum_{i \in I} u_i(\mathbf{x}^0)\mathbf{P}\mathbf{a}^{i \cdot T}(\mathbf{x}^0)$$

$$\mathbf{x}_{k+1} = \mathbf{x}^k + \alpha_{k+1}\mathbf{d}^{k+1}$$

(15) $$d^{k+1} = -\sum_{i \in I} u_i(\mathbf{x}^k)\mathbf{P}\mathbf{a}^{i \cdot T}(\mathbf{x}^k) + \frac{\left\| \sum_{i \in I} u_i(\mathbf{x}^k)\mathbf{P}\mathbf{a}^{i \cdot T}(\mathbf{x}^k) \right\|^2}{\left\| \sum_{i \in I} u_i(\mathbf{x}^{k-1})\mathbf{P}\mathbf{a}^{i \cdot L}(\mathbf{x}^{k-1}) \right\|^2} \mathbf{d}^k$$

(16) $$\alpha_{k+1} = -\frac{\sum_{i \in I} u_i(\mathbf{x}^k)\mathbf{a}^{i \cdot}(\mathbf{x}^k)\,\mathbf{d}^k}{\sum_{i \in I} u_i(\mathbf{x}^k)\,\mathbf{d}^{k+1^T}\mathbf{A}_i(\mathbf{x}^k)\mathbf{d}^{k+1}}, \quad k \geqslant 0,$$

where

$$\mathbf{A}_i(\mathbf{x}^k) = \begin{pmatrix} \frac{a_{i1}}{x_1^k} \cdot \frac{(a_{i1} - 1)}{x_1^k} & \frac{a_{i1}}{x_1^k} \cdot \frac{a_{i2}}{x_2^k} & \cdots & \frac{a_{i1}}{x_{j1}} \cdot \frac{a_{in}}{x_n^k} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{a_{in}}{x_n^k} \cdot \frac{a_{i1}}{x_1^k} & \frac{a_{in}}{x_n^k} \cdot \frac{a_{i2}}{x_2^k} & \cdots & \frac{a_{in}}{x_n^k} \cdot \frac{(a_{in} - 1)}{x_n^k} \end{pmatrix}$$

Proof. Fron (7) and Lemma 4, it follows :

$$\mathbf{x}^k = \mathbf{x}^0 + \mathbf{P}\mathbf{y}^k$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{P}(\mathbf{y}^{k+1} - \mathbf{y}^k) = \mathbf{x}^k + \alpha_{k+1}\mathbf{P}\mathbf{d}^{k+1} = \mathbf{x}^k + \alpha_{k+1}\mathbf{d}^{k+1}.$$

From (12) we have

$$\mathbf{d}^{k+1} = -\mathbf{P}\nabla^T p(\mathbf{x}^k) + \frac{\|\mathbf{P}\nabla^T p(\mathbf{x}^k)\|^2}{\|\mathbf{P}\nabla^T p(\mathbf{x}^{k-1})\|^2}\mathbf{d}^k =$$

$$= -\sum_{i \in I} u_i(\mathbf{x}^k)\mathbf{P}\mathbf{a}^{i \cdot T}(\mathbf{x}^k) + \frac{\left\| \sum_{i \in I} u_i(\mathbf{x}^k)\mathbf{a}^{i \cdot}(\mathbf{x}^k) \right\|^2}{\left\| \sum_{i \in I} u_j(x^{k-1})\mathbf{P}\mathbf{a}^{i \cdot T}(\mathbf{x}^{k-1}) \right\|^2}\mathbf{d}^k,$$

i.e. (15).

Now from (13) we get

$$\alpha_{k+1} = -\frac{\mathbf{P}\nabla^T p(\mathbf{x}^k)\,\mathbf{d}^{k+1}}{\mathbf{d}^{k+1^T}\mathbf{P}\nabla^2 p(\mathbf{x}^k)\,\mathbf{P}\mathbf{d}^{k+1}} = -\frac{\sum_{i \in I} u_i(\mathbf{x}^k)\,\mathbf{a}^{i \cdot}(\mathbf{x}^k)\,\mathbf{d}^{k+1}}{\sum_{i \in I} u_i(\mathbf{x}^k)\,\mathbf{d}^{k+1^T}\mathbf{A}_i(\mathbf{x}^k)\,\mathbf{d}^{k+1}},$$

since

$$\nabla^2 p(\mathbf{x}) = \sum_{i \in I} u_i(\mathbf{x})\mathbf{A}_i(\mathbf{x}).$$

5. Algorithm for geometric programs with linear constraints

Let us return to the program (1) in which $b_{ij} \in \mathbf{R}_+$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$. For each $\mathbf{x} \in \Omega$, consider

$$J(\mathbf{x}) = \{i : \mathbf{b}^{i \cdot}\mathbf{x} = b_i\}.$$

In what follows we assume that the following nondegeneracy condition fulfilled : with any $\mathbf{x} \in \Omega$, vectors $\mathbf{b}^{i \cdot}$, $i \in J(\mathbf{x})$ are linearly independent.

We now propose the following algorithm for solving problem (1). Starting with an arbitrary point $\mathbf{x}^0 \in \Omega$, assume that we have already constructed $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k$. To constructe $\mathbf{x}^{k+1}$ we proceed as follows: we take the set of indices $J_k = J(\mathbf{x}^k)$ and we construct projection matrix

(17) $$\mathbf{P}_k = \mathbf{E} - \mathbf{B}_{J_k}^T(\mathbf{B}_{J_k}\mathbf{B}_{J_k}^T)^{-1}\mathbf{B}_{J_k}.$$

Then calculate the quantities :

(18) $$\delta^k = \sum_{i \in I} u_i(\mathbf{x}^k)\mathbf{P}_k\mathbf{a}^{i \cdot T}(\mathbf{x}^k)$$

(19) $$\mathbf{u}_k = \sum_{i \in I} u_i(\mathbf{x}^k)(\mathbf{B}_{J_k}\mathbf{B}_{J_k}^T)^{-1}\mathbf{B}_{J_k}\mathbf{a}^{i \cdot T}(\mathbf{x}^k)$$

and test for the optimality of $\mathbf{x}^k$ (Theorem 1).

If $\delta^k \neq 0$. then we apply the method of conjugate gradients to solve the problem of minimization of $p(\mathbf{x})$ with constraints

$$\mathbf{b}^{i \cdot}\mathbf{x} - b_i = 0, \; i \in J_k.$$

However, in applying the method of conjugate gradients the following check should be made. Compute the quantity

$$\bar{\alpha}_{k+1} = \min\left\{\frac{b_i - \mathbf{b}^{i \cdot}\mathbf{x}^k}{{}_{i \cdot}\,\mathbf{d}^{k+1}} > 0 : \; i \notin J_k\right\}$$

Then

$$\mathbf{x}^{k+1} = \begin{cases} \mathbf{x}^k + \alpha_{k+1}\,\mathbf{d}^{k+1}, & \alpha_{k+1} < \bar{\alpha}_{k+1} \\ \mathbf{x}^k + \bar{\alpha}_{k+1}\,\mathbf{d}^{k+1}, & \alpha_{k+1} \geqslant \bar{\alpha}_{k+1} \end{cases}$$

In the second case, i.e. $\mathbf{x}^{k+1} = \mathbf{x}^k + \bar{\alpha}_{k+1}\mathbf{d}^{k+1}$, the process of of application of conjugate gradients stops.

If $\delta = 0$ and there exists $j \in \{1, 2, \ldots, n\}$ such that $u_j^k > 0$, construct the set of indices

$$J_k' = J_k \setminus \{j\}$$

and apply the method of conjugate gradients, corresponding to J_k'. At every step a check is made whether the point $\mathbf{x}^{k+1} \leqslant \mathbf{0}$, or not. If $\mathbf{x}^{k+1} \leqslant \mathbf{0}$, then the problem (1) has no solution.

Thus, the outlin of the algorithm for finding an optimal solution to the problem (1) is the following:

Starting from an arbitrary $\mathbf{x}^0 \in \Omega$, set $k := 0$.

Step 1. Select the set of indices $J_k = J(\mathbf{x}^k)$.

Step 2. Construct the projection matrix $\mathbf{P}_k$ as in (17).

Step 3. Calculate δ^k and $\mathbf{u}^k$ as in (18) and (19) respectively.

Step 4. If $\delta^k = 0$, go to Step 7; otherwise applying the method of conjugate gradients, find the solution $\mathbf{x}^{k+1}$ of the problem

$$\min\{p(\mathbf{x}) : \mathbf{b}\mathbf{x}^i - b_i = 0,\ i \in J_k\}.$$

Step 5. If $\mathbf{x} \leqslant \mathbf{0}$ then stop; otherwise go to Step 6. (If $\mathbf{x}^k \leqslant \mathbf{0}$ then program (1) has no optimal solution).

Step 6. Set $k := k + 1$ and go to Step 1.

Step 7. If $\mathbf{u}^k \leqslant 0$, then stop, $\mathbf{x}^k = \mathbf{x}^*$ is optimal solution of the problem (1); otherwise select $u_j^k > 0$ and the index set

$$J_k' = J_k \setminus \{j\}$$

and go to Step 2.

Remark 1. If at Step 2, $J_k = \emptyset$, then the projection matrix $\mathbf{P}_k = E$.

Remark 2. The convergence of the algorithm follows from the convergence of the method of conjugate gradients for the minimization problem of a convex function with linear constraints (see [3]).

6. Example

Consider the problem

$$p(\mathbf{x}) = x_1^{-1}\,x_2^{-1} \to \min$$

$$x_1 + x_2 \leqslant 1, \quad x_1 > 0,\ x_2 > 0.$$

Take $\mathbf{x}^0 = (3/4,\ 1/4)$.

Step 1. $J_0 = \{1\}$.

Step 2.

$$\mathbf{P}_0 = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

Step 3.

$$\delta^0 = \frac{64}{9}\begin{pmatrix} -1 \\ 1 \end{pmatrix},\ u^0 = -\frac{64}{9}$$

Step 4. As $\delta^0 \neq 0$, we find the optimal solution of the problem

$$\min\{x_1^{-1}\,x_2^{-1} : x_1 + x_2 = 1\}$$

which is $\mathbf{x}^1 = (1/2,\ 1/2)$.

Step 5. $\mathbf{x}^1 > \mathbf{0}$.

Step 1. $J_1 = \{1\}$.

Step 2.

$$\mathbf{P}_1 = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

Step 3. $\delta^1 = 0$, $u^1 = -8 < 0$.

Thus, optimal solution is $\mathbf{x}^* = (1/2,\ 1/2)$.

REFERENCES

[1] Maruşciac, I., *Programare geometrică şi aplicaţii*. Editura Dacia, Cluj-Napoca, 1978.

[2] Pchénichny B et Daniline Y., *Méthodes numériques dans les problèmes d'extremum*. Edition Mir, Moscou, 1977.

[3] Rosen, J. B., *The gradient projection method for nonlinear programming*; part I: *linear constraints*. SIAM J. Appl. Math., 8, pp. 181–217 (1960).

Received 12. IV. 1980

Universitatea „Babeş-Bolyai"
Facultatea de matematică
Str. Kogălniceanu nr. 1
3400 Cluj-Napoca, Romania