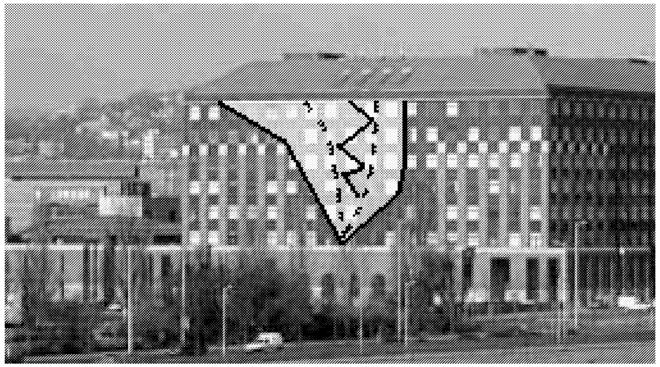
Operations Research
Report 2003-02



A sufficient optimality criteria for linearly constrained, separable concave minimization problems

Tibor Illés and Ádám B. Nagy

December 2003
Eötvös Loránd University of Sciences
Department of Operations Research

ISSN 1215 - 5918

A sufficient optimality criteria for linearly constrained, separable concave minimization problems

Tibor Illés[⋆⋆⋆] and Ádám B. Nagy[‡]

Abstract. Sufficient optimality criteria for linearly constrained, concave minimization problems is given in this paper. Our optimality criteria is based on the sensitivity analysis of the relaxed linear programming problem. Our main result is similar to that of Phillips and Rosen (1993), however our proofs are simpler and constructive.

Phillips and Rosen (1993) in their paper derived sufficient optimality criteria for a slightly different, linearly constrained, concave minimization problem using exponentially many linear programming problems. We introduced special test points and using these, for several cases, we are able to show the optimality of the current basic solution.

The sufficient optimality criteria, described in this paper, can be used as a stopping criteria for branch and bound algorithms developed for linearly constrained, concave minimization problems.

Keywords: separable concave minimization problem, linear relaxation, sensitivity analysis

Mathematics Subject Classification 2000: **90C26**, **90C20**.

⋆⋆⋆Research of Tibor Illés is supported by the Bolyai János Research Fellowship (BO/00334/00) of the Hungarian Academy of Science.

‡The author wish to acknowledge financial support received from the Hungarian Scientific Research Foundation (OTKA T038027).

1 Introduction

We consider separable concave minimization problem in the following form

$$\left.\begin{array}{c}\min \sum\limits_{j=1}^{n} f_j(x_j)\\ A\,\mathbf{x} \le \mathbf{b}\\ \mathbf{l} \le \mathbf{x} \le \mathbf{u}\end{array}\right\} \qquad (P)$$

where $A \in \mathbb{R}^{m\times n}$ is a matrix, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{l}, \mathbf{u} \in \mathbb{R}^n$ are given vectors, $\mathbf{l} \ge \mathbf{0}$, $f_j : \mathbb{R} \to \mathbb{R}$ are concave functions and $\mathbf{x} \in \mathbb{R}^n$ is the vector of the unknowns. Let us introduce the sets

$$\mathcal{A} := \{\mathbf{x} \in \mathbb{R}^n : A\,\mathbf{x} \le \mathbf{b}\} \quad \text{and} \quad \mathcal{T} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{l} \le \mathbf{x} \le \mathbf{u}\}.$$

Then the set of feasible solutions of problem (P) is defined as $\mathcal{P} = \mathcal{A} \cap \mathcal{T}$ which assume that the domain of f_j $[l_j, u_j] \subseteq \mathcal{D}f_j$ holds. Furthermore $\mathcal{P}^*$ denotes the set of optimal solutions of problem (P). If $\mathcal{P} \ne \emptyset$ then $\mathcal{P}^* \ne \emptyset$ holds, since $\mathcal{P}$ is bounded and closed. Problem (P) is one of the most simple optimization problems which do not belong to the class of convex optimization. This problem has two important theoretical properties: there is optimal solution at vertex of the polyhedron $\mathcal{P}$ [19], moreover if f_j is strict concave then each optimal solution is a vertex of the polyhedron and the problem (P) is in the class of NP-complete problems [20].

Several practical problem can be formulated by problem (P) like some control problems [1], concave knapsack problems [21], some production and transportation problems [15], production planning problems [17], process network synthesis problems [11], some network flow problems [26].

Due to the importance and applicability of model (P), the literature of possible solution methods is quite large. We know from literature three main types of algorithm: listing vertices of the polyhedron $\mathcal{P}$, cutting plane methods and branch-and-bound algorithms (BB). Several versions of BB are discussed in papers [10], [24], [22], [23], [16], [3], [6], [18], vertex enumaration procedures are used in papers [4], [8] and [9] to solve problem (P). Cutting plane algorithms are described in papers [25], [5] and [12]. There are some further methods like approximation using splines [13] or combination of BB and cutting plane algorithms [5].

Sufficient optimality criteria is given for the linearly constrained, separable concave minimization problem.

The optimality criteria is based on the linear programming relaxation of (P) and the sensitivity analysis of that linear programming problem. Our result is similar to Phillips and Rosen's result [22] but our proof is elementary and constructive, and does not require to find common solution of exponentially many linear programming problems. We introduce a test which can be efficiently used as stopping or branching criteria in a BB algorithm. The numerical implementation and testing of our BB algorithm is in progress.

Section 2 deals with the linear programming relaxation of problem (P) and the related optimality criteria known from the literature of linear programming. In section 3 using the sensitivity analysis of linear programming problem we introduce the sufficient optimality criteria for problem (P). A test point, which may violate these optimality criteria derived from the sensitivity analysis of linear programming relaxation, has been introduced in section 4. In section 4 we show, that the non existence of a violating test point for the optimal solution of the relaxed linear programming problem means that the given vertex is the optimal solution of the original problem (P).

In this paper, small (indexed) Latin (sometimes Greek) letters like x_i y_j, γ, β, $\ldots$ denote (real) numbers. Exceptions are f, f_j, $\bar{f}$, f'_k, g, g_j which are used for denoting functions and i, j, k, l are (general) indices, while m is the number of constraints in (P) and n denotes the number of variables used in the problem. Capital Latin letters likes A, B,... denote matrices, while calligraphic letters $\mathcal{A}$, $\mathcal{P}$,... denote sets. The n dimensional Euclidean space is denoted by $\mathbb{R}^n$, and the matrices of size $m \times n$, by $\mathbb{R}^{m\times n}$. Elements of the sets in $\mathbb{R}^n$ or $\mathbb{R}^m$, are unknowns of the system of linear inequalities, bounds and right hand-sides of these inequalities and columns (rows) of matrices are all vectors and has been denoted by small bold face letters $\mathbf{x}$, $\mathbf{b}$, $\mathbf{l}$, $\mathbf{u}$, $\mathbf{0}$, $\mathbf{a}_j$ etc. Furthermore, let us denote the nonnegative orthant of $\mathbb{R}^n$ by $\mathbb{R}^n_\oplus$, i.e. $\mathbb{R}^n_\oplus = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq 0\}$. The positive vectors of the Euclidean space are denoted by $\mathbb{R}^n_+$ and the indices of the unknowns by $\mathcal{J} := \{1, ..., n\}$. The summing vector of all 1's is vector $\mathbf{e}$.

2 The relaxed linear programming problem

During the solution of problem (P), for instance using a branch-and-bound algorithm, the first step is to form and solve a linear programming relaxation. Solving the linear programming relaxation and applying the post optimality analysis for the optimal solution we can introduce sufficient optimality criteria

of the original problem. In this way we can decide whether the optimal solution of the relaxed problem is the optimal solution of the original problem (P) or not.

It is necessary to introduce the following statement about properties if one dimensional concave function, Á., Császár [7], 228. page.

Proposition 2.1. Let f be one dimensional function on interval $I \subset D_f$. The following statements are equivalent

(a) *f is concave on interval I;*

(b) *let $x, y \in I$, $x \neq y$ and*

$$m(x,y) = \frac{f(y) - f(x)}{y - x}.$$

if $a, b, c \in I$, $a < b < c$ then the following holds

$$m(a,b) \geq m(a,c) \geq m(b,c);$$

(c) *for any $t \in I$, $m_t(x) = m(t,x)$ function is decreasing on $I \setminus \{t\}$;*

(a) *if $a, b, c \in I$, $a < b < c$ then*

$$m(a,b) \geq m(b,c).$$

□

The following proposition is an important consequence of the properties listed above, Á., Császár [7], 232. page.

Proposition 2.2. Let f be one dimensional concave function on open interval $I \subset D_f$, then

(a) *f is continuous on interval I;*

(b) *at any $t \in I$ the function is left and right differentiable and*

$$f'_-(t) \geq f'_+(t);$$

(c) *if $a, b, \in I$, $a < b$ then*

$$f'_+(a) \geq m(a,b) \geq f'_-(b),$$

moreover, if f is strict concave on interval I, then

$$f'_+(a) > m(a,b) > f'_-(b).$$

□

The subsections 2.1 and 2.2 summarize the relaxed linear programming problem generation pertaining to the problem (P) in order to calculate a lower bound for problem (P). Although these calculations are well known in the literature, this part is being discussed here for the sake of completeness.

In subsection 2.3 the optimality criteria of the relaxed linear programming problem are obtained.

2.1 Relaxation of the concave functions f_j on the set $\mathcal{T}$

Let us consider the linear relaxation of the concave functions $f_j : \mathbb{R} \to \mathbb{R}$ on closed interval $[l_j,\, u_j]$ as follows

$$g_j(x_j) = c_j\, x_j + d_j,$$

where

$$c_j = \frac{f_j(u_j) - f_j(l_j)}{u_j - l_j} \qquad \text{and} \qquad d_j = f_j(l_j) - \frac{f_j(u_j) - f_j(l_j)}{u_j - l_j}\, l_j = f_j(l_j) - c_j\, l_j,$$

thus

$$g_j(x_j) = c_j\, x_j + d_j = c_j\, x_j + f_j(l_j) - c_j\, l_j.$$

Then the objective function $f(\mathbf{x}) = \sum_{j=1}^{n} f_j(x_j)$ is approximated by the linear function

$$\begin{aligned} g(\mathbf{x}) &= \sum_{j=1}^{n} g_j(x_j) = \sum_{j=1}^{n} (c_j\, x_j + f_j(l_j) - c_j\, l_j) \\ &= \mathbf{c}^T\mathbf{x} + (f(\mathbf{l}) - \mathbf{c}^T\mathbf{l}) = \mathbf{c}^T\mathbf{x} + \delta \end{aligned}$$

on the set $\mathcal{P} = \mathcal{A} \cap \mathcal{T}$, where

$$\delta = f(\mathbf{l}) - \mathbf{c}^T\mathbf{l}.$$

from the properties of the function f it is easy to show that

$$f(\mathbf{x}) \geq g(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + \delta$$

holds for all $\mathbf{x} \in \mathcal{P}$.

2.2 Computing lower bound for the objective value of (P)

Lower bound for the objective value of (P) can be computed using the following linear programming problem

$$\min_{\mathbf{x} \in \mathcal{P}} \ \mathbf{c}^T\mathbf{x} + \delta \qquad (P_{LP})$$

Let us denote the optimal solution of (P_{LP}) by $\tilde{\mathbf{x}}$ then $\beta = g(\tilde{\mathbf{x}}) = \mathbf{c}^T\tilde{\mathbf{x}} + \delta$ which is the optimal objective value of it. Therefore

$$\beta = \mathbf{c}^T\tilde{\mathbf{x}} + \delta \leq f(\mathbf{x}) \leq f(\tilde{\mathbf{x}}) + (\nabla f(\tilde{\mathbf{x}}))^T (\mathbf{x} - \tilde{\mathbf{x}})$$

holds, for all $\mathbf{x} \in \mathcal{P}$, namely, a lower bound has been obtained for the optimal objective value of the problem (P). The second inequality holds because of the concavity of the function f, since the linear function

$$\bar{f}(\mathbf{x}) := f(\tilde{\mathbf{x}}) + (\nabla f(\tilde{\mathbf{x}}))^T (\mathbf{x} - \tilde{\mathbf{x}}) \qquad (1)$$

is the tangent of f in $\tilde{\mathbf{x}} \in \mathcal{P}$. [1]

2.3 Optimality criteria of the relaxed LP problem

Without loss of generality we may consider the relaxed LP problem of (P) in the following form

$$\left.\begin{array}{r} \min \mathbf{c}^T\mathbf{x} \\ A\,\mathbf{x} \leq \mathbf{b} \\ \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \end{array}\right\} \qquad (P_{LP})$$

(The constant δ, from the objective function has been deleted.) Let us denote the optimal solutions of the problem (P_{LP}) by $\mathcal{P}_c^*$.

$$\mathcal{P}_c^* = \{\mathbf{x}^* \in \mathcal{P} : \mathbf{c}^T\mathbf{x}^* \leq \mathbf{c}^T\mathbf{x},\ \mathbf{x} \in \mathcal{P}\}$$

During the sensitivity analysis we need some notations related to linear programming.

Let us denote the index set of the optimal basis by $\mathcal{J}_B \subset \mathcal{J}$, while $\mathcal{J}_N$ contains the indices of nonbasic variables. The vectors $\{\mathbf{a}_j : j \in \mathcal{J}_B\}$ are linearly independent. Obviously $\mathcal{J} = \mathcal{J}_B \cup \mathcal{J}_N$ and $\mathcal{J}_B \cap \mathcal{J}_N = \emptyset$.

[1] The condition for differentiability can be disregarded, since concave function f has subgradient at every inner point of its domain [2, 14] and so a subgradient can be considered instead of $\nabla f(\tilde{\mathbf{x}})$.

The index set $\mathcal{J}_N^l \subset \mathcal{J}_N$ ($\mathcal{J}_N^u \subset \mathcal{J}_N$) contains those indices which are at their lower (upper) bound in this basis. In our problem (P) the index set of nonbasic variables, $\mathcal{J}_N$ is partitioned into two subsets as follows $\mathcal{J}_N = \mathcal{J}_N^l \cup \mathcal{J}_N^u$.

Let $\bar{A} = [B^{-1}A]$, and $\mathbf{c}_B$ denotes the vector which contains the objective coefficients of the basic variables.

For any $\bar{\mathbf{x}} \in \mathcal{P}$ basic feasible solution of the (P_{LP}) problem the following statements are true.

$$\begin{aligned} l_i \le \bar{x}_i \le u_i \quad &\text{for all} \quad i \in \mathcal{J}_B,\\ \bar{x}_i = l_i \quad &\text{for all} \quad i \in \mathcal{J}_N^l,\\ \bar{x}_i = u_i \quad &\text{for all} \quad i \in \mathcal{J}_N^u. \end{aligned}$$

As a consequence of these, if a basic partition of the index set $\mathcal{J}$ given as $(\mathcal{J}_B, \mathcal{J}_N^l, \mathcal{J}_N^u)$ then the basic variables value can be computed as follows

$$\bar{\mathbf{x}}_B = B^{-1}\,\mathbf{b} - \sum_{j \in \mathcal{J}_N^l} l_j\,\bar{\mathbf{a}}_j - \sum_{j \in \mathcal{J}_N^u} u_j\,\bar{\mathbf{a}}_j,$$

where $\bar{\mathbf{a}}_j$ denotes the j-th column vector of matrix $\bar{A}$.

The dual problem of the (P_{LP}) has the form

$$\left.\begin{array}{c} \max\, -\mathbf{b}^T\mathbf{y} + \mathbf{l}^T\mathbf{z} - \mathbf{u}^T\mathbf{s} \\ -A^T\mathbf{y} + \mathbf{z} - \mathbf{s} = \mathbf{c} \\ \mathbf{y} \ge \mathbf{0}, \quad \mathbf{z} \ge \mathbf{0}, \quad \mathbf{s} \ge \mathbf{0} \end{array}\right\} \qquad (D_{LP}),$$

and denotes

$$\mathcal{D} = \{(\mathbf{y}, \mathbf{z}, \mathbf{s}) : -A^T\mathbf{y} + \mathbf{z} - \mathbf{s} = \mathbf{c},\, \mathbf{y} \ge \mathbf{0},\, \mathbf{z} \ge \mathbf{0},\, \mathbf{s} \ge \mathbf{0}\}$$

the set of dual feasible solutions. Let us consider the *weak duality theorem* related to (P_{LP}) and (D_{LP}) problems.

Proposition 2.3. Let $\mathbf{x} \in \mathcal{P}$ and $(\mathbf{y}, \mathbf{z}, \mathbf{s}) \in \mathcal{D}$ vectors then

$$\mathbf{c}^T\mathbf{x} \ge -\mathbf{b}^T\mathbf{y} + \mathbf{l}^T\mathbf{z} - \mathbf{u}^T\mathbf{s} \tag{2}$$

inequality holds. In (2), equality holds if and only if

$$0 = \mathbf{c}^T\mathbf{x} + \mathbf{b}^T\mathbf{y} - \mathbf{l}^T\mathbf{z} + \mathbf{u}^T\mathbf{s} = \mathbf{y}^T(\mathbf{b} - A\,\mathbf{x}) + \mathbf{z}^T(\mathbf{x} - \mathbf{l}) + \mathbf{s}^T(\mathbf{u} - \mathbf{x}). \qquad \square$$

Now we are ready to introduce the *optimality criteria* (necessary and sufficient) of the problems (P_{LP}) and (D_{LP}) as

$$\begin{array}{c} A\,\mathbf{x} \le \mathbf{b}, \quad \mathbf{l} \le \mathbf{x} \le \mathbf{u} \\ -A^T\mathbf{y} + \mathbf{z} - \mathbf{s} = \mathbf{c}, \quad \mathbf{y} \ge \mathbf{0},\, \mathbf{z} \ge \mathbf{0},\, \mathbf{s} \ge \mathbf{0} \\ \mathbf{y}\,(\mathbf{b} - A\,\mathbf{x}) = \mathbf{0}, \quad \mathbf{z}\,(\mathbf{x} - \mathbf{l}) = \mathbf{0}, \quad \mathbf{s}\,(\mathbf{u} - \mathbf{x}) = \mathbf{0}, \end{array}$$

where $\mathbf{y}\,(\mathbf{b}-A\,\mathbf{x})$, $\mathbf{z}\,(\mathbf{x}-\mathbf{l})$ and $\mathbf{s}\,(\mathbf{u}-\mathbf{x})$ denote the Hadamard (coordinatewise) product of the corresponding vectors.

Assuming that $\mathbf{x}^* \in \mathcal{P}_c^*$ is a basic solution belonging to the basis B and $\mathbf{y}^* = \mathbf{c}_B^T B^{-1} \ge \mathbf{0}$, we get that

- in case of $j \in \mathcal{J}_B$, $l_j < x_j^* < u_j$, $z_j = 0$ and $s_j = 0$ hold and thus
$$-\mathbf{a}_j^T\mathbf{y} = c_j$$
- in case of $j \in \mathcal{J}_N^l$, $l_j = x_j^*$, $z_j \ge 0$ and $s_j = 0$ hold and thus
$$z_j = c_j + \mathbf{a}_j^T\mathbf{y} \ge \mathbf{0}$$
- in case of $j \in \mathcal{J}_N^u$, $u_j = x_j^*$, $z_j = 0$ and $s_j \ge 0$ hold and thus
$$-s_j = c_j + \mathbf{a}_j^T\mathbf{y} \le \mathbf{0}.$$

Finally, we obtain a basic solution $\mathbf{x}^* \in \mathcal{P}$, which is optimal if and only if

$$\begin{aligned} \mathbf{y}^* = \mathbf{c}_B^T B^{-1} \;&\ge\; \mathbf{0} && && (3)\\ -\mathbf{c}_B^T B^{-1}\mathbf{a}_j \;&\le\; c_j, && \text{any } j \in \mathcal{J}_N^l \text{ and} && (4)\\ -\mathbf{c}_B^T B^{-1}\mathbf{a}_j \;&\ge\; c_j, && \text{any } j \in \mathcal{J}_N^u && (5) \end{aligned}$$

hold.

3 Sufficient optimality criteria

Here we formulate and prove sufficient optimality criteria for problem (P) with regard to an extremal point of set $\mathcal{P}$, concerning a basic solution.

Let us define the set $\mathcal{H} \subseteq \mathbb{R}^n$, which contains coefficients of the corresponding relaxed linear, objective functions.

The set $\mathcal{H}$ must be such that if the optimal solutions of linear programming problem related to the elements of set $\mathcal{H}$ were known then the optimal solution of (P) could be generated, too. Otherwise it would be impossible to identify the optimal solution of problem (P) using optimization method based on linear relaxation.

At first we examine $\mathcal{H}$ in general term and use the most important properties of it. Later we define sets with simple structure which approximates set $\mathcal{H}$.

The next lemma states the existence of a vector $\mathbf{h} \in \mathbb{R}^n$ for an optimal basic solution $\hat{\mathbf{x}}$ of problem (P) such that $\hat{\mathbf{x}} \in P^*_{\mathbf{h}}$ holds. Namely $\hat{\mathbf{x}}$ is element of the set of optimal solutions of the relaxed linear programming problem with objective function coefficient $\mathbf{h}$.

Lemma 3.1. *Consider problem (P). The optimal solution of (P) is denoted by $\hat{\mathbf{x}}$ and $f(\hat{\mathbf{x}}) = \min_{\mathbf{x}\in\mathcal{P}} f(\mathbf{x})$. Then*

$$\bar{f}(\hat{\mathbf{x}}) = \min_{\mathbf{x}\in\mathcal{P}} \bar{f}(\mathbf{x}),$$

where $\bar{f}(\mathbf{x}) = (\nabla f(\hat{\mathbf{x}}))^T(\mathbf{x}-\hat{\mathbf{x}})+f(\hat{\mathbf{x}})$, is a linear function defined by equation (1).

Proof. The following inequality holds because of the concavity of function f,

$$f(\mathbf{x}) \leq \bar{f}(\mathbf{x}) = (\nabla f(\hat{\mathbf{x}}))^T(\mathbf{x}-\hat{\mathbf{x}}) + f(\hat{\mathbf{x}}),$$

with strict equality at $\hat{\mathbf{x}}$, namely $f(\hat{\mathbf{x}}) = \bar{f}(\hat{\mathbf{x}})$. Consider the linear programming problem with objective function $\bar{f}(\mathbf{x})$. Then

$$f(\hat{\mathbf{x}}) = \min_{\mathbf{x}\in\mathcal{P}} f(\mathbf{x}) \leq \min_{\mathbf{x}\in\mathcal{P}} \bar{f}(\mathbf{x}) \leq \bar{f}(\hat{\mathbf{x}}) = f(\hat{\mathbf{x}})$$

from which

$$\min_{\mathbf{x}\in\mathcal{P}} \bar{f}(\mathbf{x}) = \bar{f}(\hat{\mathbf{x}})$$

is obtained. □

In this lemma, the differentiability of function f'_j on interval $[l_j, u_j]$ is used. It is easy to show that if f is not differentiable at $\bar{\mathbf{x}}$ then any inner point of the set of subgradients is also suitable for function $\bar{f}$.

It has been proved that there exists such vector $\mathbf{h}$ that the optimal solution $\hat{\mathbf{x}}$ of relaxed linear programming problem, concerning vector $\mathbf{h}$, as objective function coefficient, is also an optimal solution for the problem (P). Thus the set $\mathcal{H}$ should contain these vectors $\nabla f(\mathbf{x})$, $\mathbf{x} \in \mathcal{P}$.

For any $\bar{\mathbf{x}} \in \mathcal{P}$ basic solution, set $\mathcal{C}_B$ can be formulated as follows

$$\mathcal{C}_B = \{\mathbf{c} \in \mathbb{R}^n : \text{ vector } \mathbf{c} \text{ satisfies equation (3)–(5) }\}.$$

$\mathcal{C}_B$ contains such $\mathbf{c}$ vectors, for which $\bar{\mathbf{x}}$ is an optimal solution for linear programming problem

$$\left.\min_{\mathbf{x}\in\mathcal{P}} \mathbf{c}^T\mathbf{x}\right\} \quad (P_c).$$

Obviously, set $\mathcal{C}_B$ is not empty.

It is easy to prove the following statement about any linear programming relaxation of the problem (P).

Proposition 3.1. Consider the basic solution $\bar{\mathbf{x}} \in \mathcal{P}$, with basis B and let $\bar{\mathbf{h}} \in \mathcal{H}$ be a given vector. If $\bar{\mathbf{h}} \in \mathcal{C}_B$ then the $\bar{\mathbf{x}}$ is an optimal solution of the following linear programming problem.

$$\left.\min_{\mathbf{x}\in\mathcal{P}} \bar{\mathbf{h}}^T\mathbf{x}\right\} \quad (P_{\bar{h}}),$$

namely $\bar{\mathbf{x}} \in \mathcal{P}^*_{\bar{h}}$, where $\mathcal{P}^*_{\bar{h}}$ denotes the set of optimal solution of problem $(P_{\bar{h}})$. □

¿From this result follows that

$$\text{if} \quad \mathcal{H} \subseteq \mathcal{C}_B \quad \text{then} \quad \bar{\mathbf{x}} \in \mathcal{P}^*_h \tag{6}$$

holds for any $\mathbf{h} \in \mathcal{H}$.

We are ready to introduce and prove our main result, the sufficient optimality criteria for linearly constrained separable concave minimization problem (P).

Theorem 3.2. *Consider the linearly constrained, separable concave minimization problem (P), and suppose the functions f_j are strictly concave. Let $\bar{\mathbf{x}} \in \mathcal{P}$ be a basic solution with basis B that $\mathcal{H} \subseteq \mathcal{C}_B$ holds, then $\mathcal{P}^* = \{\bar{\mathbf{x}}\}$.*

Proof. Since $\mathcal{H} \subseteq \mathcal{C}_B$ thus $\bar{\mathbf{x}} \in \mathcal{P}^*_h$ holds for any $\mathbf{h} \in \mathcal{H}$.

There exist global minimal solution $\hat{\mathbf{x}}$ of (P) which is an extremal point of set $\mathcal{P}$. Suppose that $\hat{\mathbf{x}} \neq \bar{\mathbf{x}}$.

Let $\hat{\mathbf{h}} = \nabla f(\hat{\mathbf{x}})$. Since lemma 3.1 asserts $\hat{\mathbf{x}} \in \mathcal{P}^*_{\hat{\mathbf{h}}}$, otherwise $\bar{\mathbf{x}} \in \mathcal{P}^*_{\hat{\mathbf{h}}}$. The following relations hold,

$$f(\hat{\mathbf{x}}) = \bar{f}(\hat{\mathbf{x}}) = \bar{f}(\bar{\mathbf{x}}) > f(\bar{\mathbf{x}}), \tag{7}$$

which is a contradiction, thus $\hat{\mathbf{x}} = \bar{\mathbf{x}}$, then $\mathcal{P}^* = \{\bar{\mathbf{x}}\}$. □

The strict inequality comes from the strict concavity. If the condition of strict concavity is removed from the Theorem 3.2 then the inequality (7) will be modified

$$f(\bar{\mathbf{x}}) \geq f(\hat{\mathbf{x}}) = \bar{f}(\hat{\mathbf{x}}) = \bar{f}(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}})$$

so $f(\bar{\mathbf{x}}) = f(\hat{\mathbf{x}})$, thus $\bar{\mathbf{x}} \in \mathcal{P}^*$, but the equality $|\mathcal{P}^*| = 1$ cannot be guaranteed.

It has been proved that the sufficient optimality criteria for a basic solution $\bar{\mathbf{x}} \in \mathcal{P}$ of problem (P) with basis B is

$$\mathcal{H} \subseteq \mathcal{C}_B.$$

4 On set $\mathcal{H}$

The set $\mathcal{H}$ determines the strength of the optimality criteria. For some problems (P) and given basis, set $\mathcal{H}$ might be computed easily, however in general, the set $\mathcal{H}$ is nonlinear and nonconvex. Obviously the best approximation of the set $\mathcal{H}$ is our aim.

Set $\mathcal{H}$ contain vectors that are linear relaxation of function f at some feasible points of the problem. These vectors are closely related to the derivative of function f.

Checking optimality is equivalent to investigate the relation between two sets $\mathcal{C}_B$ and $\mathcal{H}$. Determining set $\mathcal{H}$ we have to take into account that the relation between these two sets should be examined easily.

One way to determine set $\mathcal{H}$ is to consider the range of derivative of function f on set $\mathcal{P}$. If f is strictly concave then, f' strictly decreasing, so it has inverse function g, then set

$$\mathcal{F} = \{\mathbf{y} : Ag(\mathbf{y}) = \mathbf{b} \text{ and } \mathbf{l} \le g(\mathbf{y}) \le \mathbf{u}\}$$

is the range of f' on $\mathcal{P}$, which is suitable for set $\mathcal{H}$. Set $\mathcal{F}$ can have complicated structure (nonlinear, nonconvex), so to decide whether $\mathcal{H} \subseteq \mathcal{C}_B$ is as difficult as the solution of problem (P).

If function f is quadratic then function g is linear so $\mathcal{H} = \mathcal{F}$ is also a polyhedron.

We approximate set $\mathcal{F}$, so that it is contained in a set which has less complicated structure (i.e. polyhedron).

Obviously, the determination of $\mathcal{H}$ is greatly influenced by the property of function f (strict concavity, differentiability etc.) On the other hand if the structure of set $\mathcal{H}$ is complicated (not polyhedron) then verifying the optimality criteria (6) can be very difficult. For this reason it is worth determining a set having simple structure (hyper rectangle) which encloses the set $\mathcal{H}$. If the set approximating $\mathcal{H}$ is only based on the properties of problem (P), we can get

$$\mathcal{H}_f = \{\mathbf{h} \in \mathbb{R}^n : h_j \in [f'_{j-}(u_j), f'_{j+}(l_j)]\}$$

and $\mathcal{H} \subseteq \mathcal{H}_f$ holds. Nevertheless, if we determine set approximating $\mathcal{H}$ for a given basic solution $\bar{\mathbf{x}} \in \mathcal{P}$ then let

$$\mathcal{H}_{f,\bar{\mathbf{x}}} = \{\mathbf{h} \in \mathbb{R}^n : h_j \in [c_j^l, c_j^u]\}$$

this set will contain the coefficients of all possible relaxed linear functions, where

$$c_j^u = \begin{cases} m(l_j, \bar{x}_j), & \bar{x}_j \ne l_j \\ f'_{j+}(l_j), & \text{otherwise} \end{cases} \quad \text{and} \quad c_j^l = \begin{cases} m(\bar{x}_j, u_j), & \bar{x}_j \ne u_j \\ f'_{j-}(u_j), & \text{otherwise} \end{cases}$$

It is obvious that $\mathcal{H} \subseteq \mathcal{H}_{f,\bar{\mathbf{x}}}$ holds.

¿From the Proposition 2.1. and 2.2. we can get the inequalities

$$f'_{j-}(u_j) \le c_j^l = m(\bar{x}_j, u_j) \le m(l_j, \bar{x}_j) = c_j^u \le f'_{j+}(l_j) \tag{8}$$

and therefore $\mathcal{H}_{f,\bar{\mathbf{x}}} \subseteq \mathcal{H}_f$ holds. Namely based on information from the basic solution $\bar{\mathbf{x}} \in \mathcal{P}$, a tighter set can be determined as set of coefficients for the relaxed linear objective functions approximating set $\mathcal{H}$. Philips and Rosen [22] also used the set $\mathcal{H}_{f,\bar{\mathbf{x}}}$. [2]

Now we are ready to examine the computational complexity of the decision problem, whether the relation

$$\mathcal{H}_{f,\bar{\mathbf{x}}} \subseteq \mathcal{C}_B$$

is true or false. Philips and Rosen [22] reduced this problem to the following question: whether exponentially large number of linear programming problems have a common optimal solution or not. If those linear programming relaxations of problem (P) have common optimal solution then that solution will be the optimal solution of problem (P), too. Obviously, it is enough to decide whether the extremal points of hyperrectangle $\mathcal{H}_{f,\bar{\mathbf{x}}}$ is element of $\mathcal{C}_B$ or not. This observation can save significant amount of computations, but it still needs checking whether the vertices (exponential number of points) belongs to set $\mathcal{C}_B$ or not.

Instead of following this line, we would like to choose, much more efficiently, such vertex from $\mathcal{H}_{f,\bar{\mathbf{x}}}$ which violate, at least one constraint from the set $\mathcal{C}_B$, if such point exists. We call these vertices as *test points*.

4.1 Defining a test point

Instead of checking inclusion $\mathcal{H}_{f,\bar{\mathbf{x}}} \subseteq \mathcal{C}_B$ we define a test point chosen from set $\mathcal{H}_{f,\bar{\mathbf{x}}}$, for each inequality of (3)–(5) system. Now, the coefficients of the

[2] If functions f_j are strictly concave then inequalities in (8) are strictly fulfilled.

linear objective function are the unknown variables in the inequality system (3)–(5). If there is no test point which violates at least one inequality then the inclusion $\mathcal{H}_{f,\bar{\mathbf{x}}} \subseteq \mathcal{C}_B$ should hold.

Let us define a test point, which belongs to $\mathcal{H}_{f,\bar{\mathbf{x}}}$, and violate the constraint indexed by $j \in \mathcal{J}_N^l$

$$-\mathbf{c}_B^T B^{-1}\mathbf{a}_j = -\mathbf{c}_B^T \bar{\mathbf{a}}_j \leq c_j.$$

It means, we choose such vertex of $\mathcal{H}_{f,\bar{\mathbf{x}}}$, which increase the left side of inequality and decrease the right side as much as possible. Therefore the test point $\bar{\mathbf{h}}_j$ can be defined as follows

$$\bar{h}_{ij} = \begin{cases} c_j^l, & \text{if } i = j \\ c_j^l, & \text{if } \bar{a}_{ij} > 0, i \in \mathcal{J}_B \\ c_j^u, & \text{if } \bar{a}_{ij} < 0, i \in \mathcal{J}_B \\ h_{ij}, & \text{if } , i \notin (\mathcal{J}_B \setminus \{i : \bar{a}_{ij} = 0\}) \cup \{j\}, \text{ where } h_{ij} \in [c_i^l, c_i^u]. \end{cases}$$

It is obvious that $\bar{\mathbf{h}}_j \in \mathcal{H}_{f,\bar{\mathbf{x}}}$ holds. From the construction of the test point it is clear that

$$\bar{\mathbf{h}}_B^T \bar{\mathbf{a}}_j + \bar{h}_{jj} \leq \mathbf{h}_B^T \bar{\mathbf{a}}_j + h_{jj}$$

holds for any $\mathbf{h} \in \mathcal{H}_{f,\bar{\mathbf{x}}}$, which is

$$-\bar{\mathbf{h}}_B^T \bar{\mathbf{a}}_j - \bar{h}_{jj} \geq -\mathbf{h}_B^T \bar{\mathbf{a}}_j - h_{jj}. \tag{9}$$

Now, if the test point does not violate the inequality, that is

$$0 \geq -\bar{\mathbf{h}}_B^T \bar{\mathbf{a}}_j - \bar{h}_{jj}, \tag{10}$$

then based on (9) and (10) there is no element of set $\mathcal{H}_{f,\bar{\mathbf{x}}}$ which can violate the inequality $j \in \mathcal{J}_N^l$. In general, the test point $\bar{\mathbf{h}}_k$ for any index $k \in \mathcal{J}_N^l \cup \mathcal{J}_N^u$ will be defined, using sets $\mathcal{J}_i^+$, $\mathcal{J}_i^-$, and $i \in \mathcal{J}_B$, as follows

$$\bar{h}_{ik} = \begin{cases} c_i^l, & \text{if } k \in \mathcal{J}_i^-, i \in \mathcal{J}_B \\ c_i^u, & \text{if } k \in \mathcal{J}_i^+, i \in \mathcal{J}_B \\ c_k^l, & \text{if } i = k, \text{ and } k \in \mathcal{J}_N^l \\ c_k^u, & \text{if } i = k, \text{ and } k \in \mathcal{J}_N^u \\ h_i, & i \notin (\mathcal{J}_B \setminus \{i : \bar{a}_{ik} = 0\}) \cup \{k\}, \text{ where } h_i \in [c_i^l, c_i^u] \end{cases}$$

where

$$\begin{aligned} \mathcal{J}_i^+ &= \{k \in \mathcal{J}_N^l : \bar{a}_{ik} < 0\} \cup \{k \in \mathcal{J}_N^u : \bar{a}_{ik} > 0\}, \text{ and} \\ \mathcal{J}_i^- &= \{k \in \mathcal{J}_N^l : \bar{a}_{ik} > 0\} \cup \{k \in \mathcal{J}_N^u : \bar{a}_{ik} < 0\}. \end{aligned}$$

Based on these observations, we can get the following proposition.

Proposition 4.1. If test point $\bar{\mathbf{h}}_k$ does not violate the inequality $k \in \mathcal{J}_N^l \cup \mathcal{J}_N^u$ then no point $\mathbf{h} \in \mathcal{H}_{f,\bar{\mathbf{x}}}$ violates either. □

Moreover, in case of $j \in \mathcal{J}_N^l$ ($j \in \mathcal{J}_N^u$)

$$-\bar{\mathbf{h}}_{B,j}^T \bar{\mathbf{a}}_j > c_j^l \qquad (-\bar{\mathbf{h}}^T{}_{B,j} \bar{\mathbf{a}}_j < c_j^u)$$

the test point $\bar{\mathbf{h}}_j$ violates the optimality criteria which belongs to the variable j.

We can determine a test point for testing inequality system

$$\bar{\mathbf{h}}_B^T B^{-1} \geq \mathbf{0}.$$

Let matrix $\bar{B} = B^{-1}$ and let $\bar{\mathbf{b}}_i$ denote the i-th column of matrix $\bar{B}$, then

$$\bar{h}_{ji} = \begin{cases} c_i^l, & \text{if } b_{ji} > 0, \ j \in \mathcal{J}_B \\ c_i^u, & \text{if } b_{ji} < 0, \ j \in \mathcal{J}_B \\ h_i, & \text{if } j \in \mathcal{J}_N^l \cup \mathcal{J}_N^u \cup \{j \in \mathcal{J}_B : \bar{b}_{ji} = 0\} \text{ where, } h_i \in [c_j^l, c_j^u] \end{cases}$$

In this case $\bar{\mathbf{h}}_{j,B}^T \bar{\mathbf{b}}_i \geq 0$ holds, and any vector $\mathbf{h} \in \mathcal{H}_{f,\bar{\mathbf{x}}}$ satisfies the i-th nonnegativity condition.

Therefore, instead of testing 2^n vertices of hyperrectangle $\mathcal{H}_{f,\bar{\mathbf{x}}}$, it is enough to determine n test points in order to check whether the inclusion $\mathcal{H}_{f,\bar{\mathbf{x}}} \subseteq \mathcal{C}_B$ holds or not.

Let us introduce the index set $\mathcal{K}$

$$\mathcal{K} = \{i : \bar{\mathbf{h}}_i \text{ test point violates } i\text{-th inequality}\}.$$

It is obvious that, the equality $\mathcal{K} = \emptyset$ leads to $\mathcal{H}_{f,\bar{\mathbf{x}}} \subseteq \mathcal{C}_B$, thus $\bar{\mathbf{x}} \in \mathcal{P}^*$ holds. The decision, whether basis $\bar{\mathbf{x}} \in \mathcal{P}$ is optimal solution for problem (P), can be performed as follows

1. generate set $\mathcal{H}_{f,\bar{\mathbf{x}}}$,
2. using matrices B^{-1} and $B^{-1}A_N$ generate test point $\bar{\mathbf{h}}_j$,
3. perform the checking of the test points, if there is no index j for which test point $\bar{\mathbf{h}}_j$ violates j-th condition then $\bar{\mathbf{x}}_j$ is optimal solution for problem (P).

Nevertheless, if any test point $\bar{\mathbf{h}}_j$ can be founded which violates j-th condition, can we conclude that $\bar{\mathbf{x}} \in \mathcal{P}$ is not an optimal solution of the problem (P)? Unfortunately we can not, because we do not know, how good is the approximation of set $\mathcal{H}$ by set $\mathcal{H}_{f,\bar{\mathbf{x}}}$.

Since sets $\mathcal{H}$ and $\mathcal{H}_{f,\bar{\mathbf{x}}}$ significantly depend on bounds l_j and u_j, thus it is expected that if the diameter of set $\mathcal{H}_{f,\bar{\mathbf{x}}}$ is rapidly decrease then a branch–and–bound type algorithms is efficient for solving problem (P).

In case of branch–and–bound type algorithms the partition is to be performed in a way that the diameter of the set $\mathcal{H}_{f,\bar{\mathbf{x}}}$ decreases. The numerical testing of the branch–and–bound algorithm, using test point introduced above, is in progress.

References

[1] Apkarian, P. and Tuan, H. D. Concave programming in control theory, *J. Global Optim.*, 15: 343–370, 1999.

[2] Bazaraa, M. S., Sherali, H. D. and Shetty, C. M., *Nonlinear Programming: Theory and Algorithms* John Wiley & Sons., Inc., New York, 1993.

[3] Benson, H. P. Separable concave minimization via partial outer approximation and branch and bound, *Oper. Res. Lett.*, 9: 389–394, 1990.

[4] Benson, H. P. and Sayin, S. A finite concave minimization algorithm using branch and bound and neighbor generation, *J. Global Optim.*, 5: 1–14, 1994.

[5] Bretthauer, K. M. and Cabot, A. V., A composite branch and bound, cutting plane algorithm for concave minimization over a polyhedron, *Comput. Oper. Res.*, 21: 777–785, 1994.

[6] Cabot, A. V. and Erenguc, S. S. A branch and bound algorithm for solving a class of nonlinear integer programming problems, *Naval Res. Logist. Quart.*, 33: 559–567, 1986.

[7] Császár Á., *Valós analízis I.*, Tankönyvkiadó, Budapest, 1983.

[8] Dyer, M. E. The complexity of vertex enumeration methods *Math. Oper. Res.*, 8: 381–402, 1983.

[9] Dyer, M. E. and Proll, L. G. An algorithm for determining all extreme points of a convex polytope, *Math. Programming*, 12: 81–96, 1977.

[10] Falk, J. E. and Soland, R. M. An Algorithm for Separable Nonconvex Programming Problems. *Management Science*, 15: 550–569, 1969.

[11] Friedler, F. and Fan, L. T. and Imreh, B. Process network synthesis: problem definition, *Networks*, 31: 119–124, 1998.

[12] Hoffman, K. L. A method for globally minimizing concave functions over convex sets, *Math. Programming*, 20: 22–32, 1981.

[13] Kontogiorgis, S. Practical piecewise-linear approximation for monotropic optimization, *INFORMS J. Comput.*, 12: 324–340, 2000.

[14] Kovács M., *A nemlineáris programozás elmélete*, Typotex, Budapest, 1997.

[15] Kuno, T. and Utsunomiya, T. Lagrangian based branch-and-bound algorithm for production-transportation problems, *J. Global Optim.*, 18: 59–73, 2000.

[16] Lamar, B. W. Nonconvex optimization over a polytope using generalized capacity improvement, *J. Global Optim.*, 7: 127–142, 1995.

[17] Liu, Ming-Long and Sahinidis, N. V. and Shectman, J. P. Planning of chemical process networks via global concave minimization, *Global optimization in engineering design*, Kluwer Acad. Publ. 1996. Dordrecht, 195–230.

[18] Locatelli, M. and Thoai, N. V., Finite exact branch-and-bound algorithms for concave minimization over polytopes, *J. Global Optim.*, 18: 107–128, 2000.

[19] Luenberger, D. G. Intoduction to linear and nonlinear programming, *Addison-Wesley Publishing Co.*, Massachusetts, 1973.

[20] Murty, K. G. and Kabadi, S. N. Some NP-complete problems in quadratic and nonlinear programming, *Math. Programming*, 39: 117–129, 1987.

[21] Moré, J. J. and Vavasis, S. A. On the solution of concave knapsack problems, *Math. Programming*, 49: 397–411, 1990/1991.

[22] Phillips, A. T. and Rosen, J. B. Sufficient conditions for solving linearly constrained separable concave global minimization problems, *J. Global Optim.*, 3: 79–94, 1993.

[23] Rosen, J. B. Global minimization of a linearly constrained concave function by partition of feasible domain, *Math. Oper. Res.*, 2: 215–230, 1983.

[24] Shectman, J. P. and Sahinidis, N. V. A finite algorithm for global minimization of separable concave programs, *J. Global Optim.*, 12: 1–35, 1998.

[25] Tuy, H. and Thieu, T. V. and Thai, Ng. Q., A conical algorithm for globally minimizing a concave function over a closed convex set, *Math. Oper. Res.*, 10: 498–514, 1985.

[26] Yajima, Y. and Konno, H. An algorithm for a concave production cost network flow problem, *Japan J. Indust. Appl. Math.*, 16: 243–256, 1999.

Tibor Illés
Department of Operations Research
Eötvös Loránd University of Sciences
Pázmány Péter sétány 1/C.
H-1117 Budapest
E-mail: illes@math.elte.hu

Ádám B. Nagy
Department of Computer Science
University of Veszprém
Egyetem u. 10.
H-8200 Veszprém
E-mail: nagy@dcs.vein.hu

Earlier Research Reports

1991-01 T. Illés, J. Mayer and T. Terlaky: A new approach to the colour matching problem

1991-02 E. Klaszky, J. Mayer and T. Terlaky: A geometric programming approach to the chanel capacity problem

1992-01 Edvi T.: Karmarkar projektív skálázási algoritmusa

1992-02 Kassay G.: Minimax tételek és alkalmazásaik

1992-03 T. Illés, I. Joó and G. Kassay: On a nonconvex Farkas theorem and its applications in optimization theory

1992-04 Interior point methods. Proceedings of the IPM 93. Workshop jan. 5. 1993

Recent Operations Research Reports

2003-01 Zsolt Csizmadia and Tibor Illés: New criss-cross type algorithms for linear complementarity problems with sufficient matrices

2003-02 Tibor Illés and Ádám B. Nagy: A sufficient optimality criteria for linearly constrained, separable concave minimization problems