

A NEW BRANCH AND BOUND APPROACH FOR CONCAVE MINIMIZATION PROBLEMS

Reiner Horst

Fachbereich Mathematik der Technischen Hochschule D 61 Darmstadt,
Schloßgartenstr. 1

Introduction

In this paper we shall consider the following programming problem:
Find a vector $\bar{x} \in D$ satisfying

$$f(\bar{x}) = \min_{x \in D} f(x). \quad (\text{Problem P})$$

It is assumed throughout this paper that the set $D \subset \mathbb{R}^n$ is convex, compact and $f: A \rightarrow \mathbb{R}$ is a concave function on an open convex set A , $D \subset A \subset \mathbb{R}^n$.

For solving problem P a branch and bound algorithm is proposed, which solves a sequence of problems in each of which the objective function is linear. The main difference between this approach and previous approaches ([1], [4]) is the use of simplex-partitions instead of rectangular ones and a different refining rule such that the convergence proof does not rely on the concept of convex envelopes and the algorithm handles non-separable functions.

In the next section we shall give some useful preparatory results. The subsequent section contains a description of the algorithm and the last one an associated convergence theorem.

A generalization of this algorithm will appear in [3]. Proofs of statements given in [3] will be omitted here.

Simplex-partitions and affine subfunctions

Starting with a n -dimensional enclosing simplex S^1 , $D \subset S^1 \subset A$, we shall get a sequence of subproblems by partitioning S^1 in from step to step refined subsimplices S^{ki} and by constructing affine underestimating functions $L^{ki}: S^{ki} \rightarrow \mathbb{R}$. This will be done in such a way that each L^{ki} will be the best convex subfunction of f over S^{ki} , that is for all convex functions $g^{ki}: S^{ki} \rightarrow \mathbb{R}$ satisfying $g^{ki}(x) \leq f(x) \forall x \in S^{ki}$ we shall have $g^{ki}(x) \leq L^{ki}(x) \leq f(x) \forall x \in S^{ki}$. The chosen refining rule than will

yield a convergence theorem.

Definition 1:

Let S be a n -dimensional simplex in the euclidian space \mathbb{R}^n . A set $\{S^1, \dots, S^r\}_S$ of finitely many n -dimensional simplices S^i ($i=1, \dots, r$) is said to be a simplex-partition of S if

$$S = \bigcup_{i=1}^r S^i,$$

$$S^i \cap S^j = S^{(ij)}, \quad i, j \in \{1, \dots, r\}, \quad i \neq j,$$

where $S^{(ij)}$ denotes a $(n-1)$ -dimensional simplex-face of both S^i and S^j .

Lemma 1:

Let $S \subset \mathbb{R}^n$ be a n -dimensional simplex (n -simplex), characterized by its $n+1$ affinely independent vertices $x_1, \dots, x_{n+1} \in \mathbb{R}^n$. We write $S = \{x_1, \dots, x_{n+1}\}$. Moreover let x be a point in S but no vertex:

$$x = \sum_{i=1}^{n+1} \kappa_i x_i, \quad \kappa_i \geq 0 \quad (i=1, \dots, n+1), \quad \sum_{i=1}^{n+1} \kappa_i = 1;$$

$x \neq x_i$ ($i=1, \dots, n+1$). Replacing one vertex x_i of S with $\kappa_i > 0$ by x we obtain a n -subsimplex $S^i = \{x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_{n+1}\}$ of S .

The set of all by means of x in this way constructible subsimplices form a simplex-partition of S .

Proof:

It is well known from linear algebra, that given a set of affinely independent vectors $S = \{x_1, \dots, x_{n+1}\}$ and a vector $x = \sum_{i=1}^{n+1} \kappa_i x_i, \kappa_i \geq 0$ ($i=1, \dots, n+1$), $\sum_{i=1}^{n+1} \kappa_i = 1$; $x \neq x_i$ ($i=1, \dots, n+1$), then if any vector $x_i \in S$ for which $\kappa_i > 0$ is removed from S and x is added to S , the new collection S^i of $(n+1)$ vectors is also a set of affinely independent vectors. It follows that all S^i constructed according to Lemma 1 are n -dimensional subsimplices of S and obviously we have $\bigcup S_i \subset S$. To prove $S \subset \bigcup S^i$ take any element $\bar{x} \in S$, $\bar{x} \neq x$, and consider the ray emanating from x through \bar{x} up to the boundary point $x^* \in S$ where it leaves S ($x^* = x$ is possible):

$$x^* = \sum_{j=i_1}^{i_m} \kappa_j^* x_j, \quad \kappa_j^* \geq 0 \quad (j=i_1, \dots, i_m), \quad \sum_{j=i_1}^{i_m} \kappa_j^* = 1;$$

$m < n+1$ since x^* is element of a simplex-face of dimension less than n .
We have

$$\bar{x} = \beta x + (1-\beta)x^*, \quad 0 \leq \beta < 1$$

or

$$\bar{x} = \beta x + \sum_{j=i_1}^m \beta_j x_j, \quad \beta_j = (1-\beta)K_j^x,$$

$$\beta_j \geq 0 (j=i_1, \dots, i_m), \quad \beta + \sum_{j=i_1}^m \beta_j = 1$$

proving $x \in \cup S^i$.

$S^i \cap S^j = S^{(ij)}$ ($i \neq j$), where $S^{(ij)}$ is a $(n-1)$ -dimensional simplex-face of both S^i and S^j , results immediately from the construction of S^i and S^j .

Lemma 2:

Let $S = \{x_1, \dots, x_{m+1}\} \subset \mathbb{R}^n$ be a n -simplex and $f: S \rightarrow \mathbb{R}$ be a concave function. Then there is exactly one affine function $L: \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying $L(x_i) = f(x_i)$ ($i=1, \dots, n+1$). Moreover $L(x)$ is the convex envelope of f over S .

Proof:

Every affine function L may be represented by $L(x) = \lambda^T x + \lambda_0$, $\lambda \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}$. The determinant of the system of linear equations

$$\lambda^T x_i + \lambda_0 = f(x_i) \quad (i=1, \dots, n+1)$$

is different from zero, since the $n+1$ vertices x_i ($i=1, \dots, n+1$) are affinely independent. $L(x) \leq f(x)$ for all $x \in S$ and moreover that $L(x)$ actually is the convex envelope of f over S (the best convex subfunction in the sense mentioned above) follows easily by concavity of f . A rigorous proof can be found in [2].

The minimum values of f and L over S are attained in a vertex of S .
It follows

$$\min_{x \in S} f(x) = \min_{x \in S} L(x) .$$

Minimizing L over S instead of f yields the minimum value of f . Unfortunately this is not true for minimizing points:

$$\{\bar{x} \mid f(\bar{x}) = \min_{x \in S} f(x)\} \subset \{\bar{x} \mid L(\bar{x}) = \min_{x \in S} L(x)\}$$

and in general there are minimizing points of L not minimizing f over S .

In our approach for solving Problem P we shall solve a sequence of sub-problems in each of which we minimize L over $S \cap D$ instead of f .

An associated convergence theorem stated in the last section will show that by suitable refining of simplex-partitions it is possible to generate a sequence of points $\{x^k\}$ having accumulation points, each of which yields the desired solution value. But the above discussion explains why we will get weaker results for the solutions points.

The next two Lemmata state how to choose a suitable refining rule. We omit the proofs, here because they are given in [3].

Lemma 3:

Let $S^1 \subset \mathbb{R}^n$ be a n -simplex and construct a sequence of n -simplices in the following way: $S^k \rightarrow S^{k+1}$ ($k=1,2,\dots$): Choose the midpoint x^k of one of the longest edges of S^k and let S^{k+1} be one of the two subsimplices constructed according to Lemma 1 using x^k instead of x . Then there is a $\tilde{x} \in \mathbb{R}^n$ satisfying

$$\lim_{k \rightarrow \infty} S^k := \bigcap_{k=1}^{\infty} S^k = \{\tilde{x}\}.$$

Lemma 4:

Let $S^1 \subset \mathbb{R}^n$ be a n -simplex and $\{S^k\}_{k=1}^{\infty}$ be a sequence of n -simplices constructed from S^1 according to Lemma 3. Furthermore let $f: A \rightarrow \mathbb{R}$ be a concave function on an open set A , $S^1 \subset A \subset \mathbb{R}^n$, and $\{L^k\}_{k=1}^{\infty}$ the sequence of affine functions each L^k defined over S^k according to Lemma 2. Then we have

- a) $L^{k-1}(x) \leq L^k(x) \leq f(x)$ for all $x \in S^k$, $k=1,2,\dots$;
- b) the sequence $\{L^k\}$ is equicontinuous in the following sense:
For every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon, x^*, \varepsilon)$ not dependent on k such that

$$|L^k(x) - L^k(x^*)| \leq \varepsilon$$

whenever x^* , $x \in S^k$ and $|x-x^*| \leq \delta$.

Description of the AlgorithmStep 0:

0.1.: Choose a n -simplex $S^1 \subset \mathbb{R}^n$ such that $D \subset S^1 \subset A$.

(A is supposed to be "large" enough of course such that S^1 exists.)

(If in problem (P) $D = D_1 \cap D_2$, D_1 convex, compact and

$D_2 := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \geq 0, i=1, \dots, n\}$, then one method to construct S^1 is solving the optimization problem

$$z_0 := \max_{x \in D} \sum_{i=1}^n x_i$$

and taking $S^1 = \{x_1^1, \dots, x_{n+1}^1\}$, $x_{n+1}^1 = \ominus$, the origin, $z_i^1 = z_0 e_i$ ($i=1, \dots, n$), where e_i denotes the i -th unit-coordinate vector.)

0.2.: Determine the affine function $L^0(x)$ over S^1 according to Lemma 2.

0.3.: Compute x^0 from

$$L^0(x^0) := \min_{x \in D} L^0(x) .$$

0.4.: Compute $f(x^0)$. If $L^0(x^0) = f(x^0)$, then stop the algorithm. If $L^0(x^0) < f(x^0)$, then go to step 1.

Step 1:

1.1.: Construct the simplex-partition $\{S^{11}, S^{12}\}_{S^1}$ according to Lemma 3.

1.2.: Determine the affine functions $L^{11}: S^{11} \rightarrow \mathbb{R}$, $L^{12}: S^{12} \rightarrow \mathbb{R}$ according to Lemma 2.

1.3.: Compute x^{1i} from

$$L^{1i}(x^{1i}) = \min_{x \in D \cap S^{1i}} L^{1i}(x)$$

for those $i \in \{1, 2\}$ with $D \cap S^{1i} \neq \emptyset$.

1.4.: Compute x^1 from

$$L^1(x^1) := \min_i L^{1i}(x^{1i}).$$

1.5.: Compute $f(x^1)$, If $L^1(x^1) = f(x^1)$, then stop the algorithm. If $L^1(x^1) < f(x^1)$, then go to step 2.

Step k (k=2,3,...):

Assume $x^{k-1} \in S^{k-1,k}$.

k.1.: Construct the simplex-partition $\{S^{k1}, S^{k2}\}_{S^{k-1,k}}$ of $S^{k-1,k}$ according to Lemma 3.

k.2.: Determine the affine functions $L^{ki}: S^{ki} \rightarrow \mathbb{R}$ (i=1,2) according to Lemma 2.

k.3.: Set $S^{k,2+i} = S^{k-1,i}$ (i=1,...,k-1),
 $L^{k,2+i} = L^{k-1,i}$ (i=1,...,k-1).

k.4.: Compute x^{ki} from

$$L^{ki}(x^{ki}) = \min_{x \in D \cap S^{ki}} L^{ki}(x), \quad D \cap S^{ki} \neq \emptyset \quad (i=1, \dots, k+1)$$

(We have to take into account all subsimplices of S constructed up to now and look for the minimum over the collection of all affine subfunctional at hand. x^{ki} is already known for $i > 2$).

k.5.: Compute x^k from

$$L^k(x^k) := \min_i L^{ki}(x^{ki}).$$

k.6.: Compute $f(x^k)$. If $L^k(x^k) = f(x^k)$, then stop the algorithm. If $L^k(x^k) < f(x^k)$, then go to step k+1.

Main properties of the algorithm

The main properties of the algorithm are stated in the following theorem.

Theorem

- a) Let $\bar{x} \in D$ be a solution of problem P. Then
 $L^k(x^k) \leq L^{k+1}(x^{k+1}) \leq f(\bar{x}) \leq \min_{j=1, \dots, k+1} f(x^j) \leq \min_{j=1, \dots, k} f(x^j) \quad k=1, 2, \dots;$
- b) If the algorithm stops after step j, then x^j is a solution of problem P;

c) If the algorithm is not terminated after finitely many steps, then it generates a sequence $\{x^k\}$ of feasible points one accumulation point of which is a solution. Moreover the limit of the numbers $\{L^k(y^k)\}$ is $f(\bar{x})$ where $\{y^k\}$ is any convergent subsequence of $\{x^k\}$.

The theorem is a consequence of Lemmata 1 to 4, a more general version is proved in [3]. The simple illustrative example discussed in the course of the presentation at the conference is also given in [3].

First feq tests had been rather successful finding good upper and lower bounds for the solution after a few steps.

References

- [1] J.E. Falk and R.M. Soland, An algorithm for separable nonconvex programming problems, Management Sci. 15(1969), 550-569.
- [2] R. Horst, Zur Charakterisierung affin-linearer Hüllfunktionale, (to appear in Zeitschrift für Angewandte Mathematik und Mechanik)
- [3] R. Horst, An algorithm for nonconvex programming problems, (to appear in Mathematical Programming)
- [4] R.M. Soland, An algorithm for separable nonconvex programming II: nonconvex constraints, Management Sci. 17(1971), 759-773.