<table>
<thead>
<tr>
<th>Title</th>
<th>On subdivision strategies in the conical algorithm for concave minimization (The advances and applications of optimization method)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Ishihama, Tomohiro; Kuno, Takahito</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (2012), 1773: 177-185</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2012-01</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/171702">http://hdl.handle.net/2433/171702</a></td>
</tr>
<tr>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
On subdivision strategies in the conical algorithm for concave minimization

Tomohiro Ishihama, and Takahito Kuno
石濱友裕, 久野誉人
Department of Computer Science, University of Tsukuba
筑波大学コンピュータサイエンス専攻

Abstract. We present a new proof of the convergence of the conical algorithm for concave minimization under a pure $\omega$-subdivision strategy. For this purpose, we introduce a weaker condition of nondegeneracy for sequences of nested cones generated in the algorithm. We show that this condition is not only useful for proving the convergence but also suggests a possible class of convergent subdivision strategies.

Keywords: Global optimization, concave minimization, conical algorithm, $\omega$-subdivision.

1 Introduction

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a quasiconcave function and consider the following problem:

$$
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad Ax \leq b, \quad x \geq 0,
\end{align*}
$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Over the past four decades, many algorithms have been developed to solve this multiextremal global optimization problem. Among them is the conical algorithm, the convergence of which is the main concern in this paper. This algorithm uses a branch-and-bound technique, and repeatedly deletes a portion of the feasible set intersected with some polyhedral cone unless it contains some $x$ such that $f(x) < \gamma$ for the incumbent value $\gamma$. To subdivide the cone for branching, Tuy proposed the concept of $\omega$-subdivision process in 1964 [9]. In this subdivision process, each cone is subdivided radially from a feasible point, which is given as a byproduct of the bounding operation. In spite of extensive studies, the convergence of this process remained an open question until the late 1990s. In 1991, Tuy [10] showed that the convergence is guaranteed if the $\omega$-subdivision process satisfies a certain kind of nondegeneracy condition. To the present, however, yet none has succeeded in proving the nondegeneracy of the $\omega$-subdivision process. At last, after ten years of [10], Jaumard and Meyer showed the convergence of $\omega$-subdivision with no help of the nondegeneracy [3, 5]. Around the same time, Locatelli separately proved it in a different way, but still without using the nondegeneracy [7].

Those earlier studies allowed us to apply the algorithm with $\omega$-subdivision to (1) without having to worry about its convergence. However, it remains unsolved whether or not the $\omega$-subdivision process is nondegenerate. In this paper, we introduce a new notion of nondegeneracy, called pseudo-nondegeneracy, which is a weaker condition than the original nondegeneracy. We then show the $\omega$-subdivision process is pseudo-nondegenerate, and give another proof of the
convergence using the pseudo-nondegeneracy, along the lines of Tuy's research [10]. Furthermore, we present a certain class of pseudo-nondegenerate subdivision strategies, which includes $\omega$-subdivision as a special case.

2 Conical Algorithm

Let us denote the feasible set of (1) by the intersection of two polyhedral sets:

$$D = \{x \in \mathbb{R}^n \mid Ax \leq b\}, \quad \Lambda_1 = \{x \in \mathbb{R}^n \mid x \geq 0\}.$$ 

For simplicity, we assume that

(a) the feasible set $D \cap \Lambda_1$ is bounded, and
(b) the origin $0 \in \mathbb{R}^n$ is a vertex of $D \cap \Lambda_1$ and incident to exactly $n$ linearly independent edges.

Also let $\gamma$ be an arbitrary number satisfying

$$\gamma < f(0),$$

and let

$$C_{\gamma} = \{x \in \mathbb{R}^n \mid f(x) \geq \gamma\}.$$ 

In addition to (a) and (b), assume through the paper that

(c) $C_{\gamma}$ is a bounded set.

In the rest of this section, after providing some basic operations needed in the conical algorithm, we give a description of the algorithm.

$\gamma$-extension

For any nonzero vector $d \in \mathbb{R}^n$, let

$$\theta = \max\{\alpha \mid f(\alpha d) \geq \gamma, \alpha d \in C_{\gamma}, \alpha \geq 0\}.$$

We refer to $q = \theta d$ as the $\gamma$-extension along $d$. Note that $\|q\|$ is always finite under the assumption (c).

Deletion test

Suppose that $q_1, \ldots, q_n$ are $\gamma$-extensions and linearly independent. Let $\Lambda$ denote the simplicial cone spanned by $q_1, \ldots, q_n$, i.e.,

$$\Lambda = \text{con}Q \equiv \{x \in \mathbb{R}^n \mid x = \sum_{j=1}^{n} q_j \lambda_j, \lambda \geq 0\},$$

where

$$Q = [q_1, \ldots, q_n] \in \mathbb{R}^{n \times n}.$$
Also let
\[ G^{-} = \{ x \in \mathbb{R}^n \mid eQ^{-1}x \leq 1 \}, \]
where \( e \in \mathbb{R}^n \) is the all-ones row vector. Note that
\[ 0 \in G^{-} \cap \Lambda \subset C_{\gamma}. \tag{2} \]
We can check if \( D \cap \Lambda \subset G^{-} \cap \Lambda \), by solving a linear programming problem:
\[
\begin{align*}
\text{maximize} & \quad eQ^{-1}x \\
\text{subject to} & \quad Ax \leq b, \quad Q^{-1}x \geq 0.
\end{align*}
\tag{3}
\]
Let \( \omega(\Lambda) \) be an optimal solution of (3) and \( \zeta(\Lambda) \) the optimal value. If \( \zeta(\Lambda) \leq 1 \), then \( D \cap \Lambda \subset G^{-} \cap \Lambda \), and hence \( f(x) \geq \gamma \) for any \( x \in D \cap \Lambda \). If \( \gamma = f(x^*) \) for the incumbent \( x^* \), we can remove \( \Lambda \) from consideration because it contains no feasible solution better than \( x^* \).

**\( \omega \)-subdivision**

Starting from \( \Lambda = \Lambda_1 \), the conical algorithm subdivides \( \Lambda \) recursively, via a given point \( p \in \Lambda \). Let \( J \) be an index set such that \( i \in J \) if \( p \) is linearly independent of \( q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n \). Let \( \text{ext}(p) \) denote the \( \gamma \)-extension along \( p \). Then \( \Lambda \) is partitioned into \( |J| \) children
\[ \text{con}[q_1, \ldots, q_{i-1}, \text{ext}(p), q_{i+1}, \ldots, q_n], \quad j \in J. \]
Recall that the optimal solution \( \omega(\Lambda) \) of (3) is a point in \( \Lambda \). If we choose \( p = \omega(\Lambda) \), this subdivision is called **\( \omega \)-subdivision**.

**Algorithm description**

For a given tolerance \( \epsilon \geq 0 \), the conical algorithm can be described as follows:

**Conical Algorithm with \( \omega \)-subdivision**

**Step 1. (Initialization)** Let \( x^* \leftarrow 0, f^* \leftarrow f(0), \gamma \leftarrow f^* - \epsilon \). Solve (3) with \( \Lambda = \Lambda_1 \) to obtain \( \omega(\Lambda_1) \) and \( \zeta(\Lambda_1) \). If \( \zeta(\Lambda_1) \leq 1 \), then terminate. Otherwise, let \( k \leftarrow 1 \), and \( \mathcal{K} \leftarrow \{ \Lambda_1 \} \).

**Step 2. (Subdivision)** Select a \( \Lambda_k \in \arg \max \{ \zeta(\Lambda) \mid \Lambda \in \mathcal{K} \} \). Let \( \omega^k \leftarrow \omega(\Lambda_k), \zeta^k \leftarrow \zeta(\Lambda_k) \), and subdivide \( \Lambda_k \) via \( \omega^k \). Let \( \mathcal{L} \) denote the set of the resulting subcones.

**Step 3. (Deletion test)** For each cone \( \Lambda \in \mathcal{L} \), solve (P) to obtain \( \omega(\Lambda) \) and \( \zeta(\Lambda) \). If \( \zeta(\Lambda) > 1 \), then add \( \Lambda \) to \( \mathcal{K} \).

**Step 4. (Updating the incumbent)** If \( f(\omega(\Lambda)) < f^* \) for some \( \Lambda \in \mathcal{L} \), then \( x^* \leftarrow \omega(\Lambda), f^* \leftarrow f(\omega(\Lambda)), \) and \( \gamma \leftarrow f^* - \epsilon. \)

**Step 5. (Optimality test)** Let \( \mathcal{K} \leftarrow \mathcal{K} \setminus \{ \Lambda_k \} \). If \( \mathcal{K} \) is empty, terminate. Otherwise, return to Step 2 with \( k \leftarrow k + 1 \).

If this algorithm terminates after finitely many iterations, then \( x^* \) is a globally \( \epsilon \)-optimal solution of (1), i.e., it holds that
\[ f^* = f(x^*) \leq f(x) + \epsilon, \quad \forall x \in D \cap \Lambda_1. \]
3 Pseudo-Nondegeneracy

Suppose the conical algorithm does not terminate and generates an infinite sequence of nested cones:

\[ \Lambda_1 \supset \cdots \supset \Lambda_k \supset \Lambda_{k+1} \supset \cdots, \]

where \( \Lambda_{k+1} \) is a cone generated by subdividing \( \Lambda_k \) via \( \omega^k \). For each \( k \), the cone \( \Lambda_k \) is spanned by \( n \) linearly independent vectors \( q_k^j \), which are \( \gamma \)-extensions and the columns of \( Q_k \). Let us denote the problem (3) with \( \Lambda = \Lambda_k \) by

\[
\begin{align*}
(P_k) & \quad \text{maximize} \quad eQ_k^{-1}x \\
& \text{subject to} \quad Ax \leq b, \quad Q_k^{-1}x \geq 0.
\end{align*}
\]

Note that we can assume

\[ \zeta^k > 1, \quad k = 1, 2, \ldots \] (4)

Otherwise, \( \Lambda_k \) must have been discarded by the deletion test.

Let us denote by \( y^k \) the intersection point of the ray emanating from \( 0 \) to \( \omega^k \) with the boundary of \( G_k^{-} \).

**Definition 1.** [10] The sequence of nested cones \( \{ \Lambda_k \mid k = 1, 2, \ldots \} \) is said to be nondegenerate if there exists a subsequence \( \{ k_t \mid t = 1, 2, \ldots \} \) and constant \( M \) such that

\[ \|eQ_{k_t}^{-1}\| \leq M, \quad t = 1, 2, \ldots. \]

Also, the subdivision process is nondegenerate if every sequence of nested cones is nondegenerate.

**Proposition 1.** [10] If \( \{ \Lambda_k \mid k = 1, \ldots \} \) is nondegenerate, then

\[
\lim_{k \to +\infty} \inf \| \text{ext}(\omega^k) - y^k \| = 0.
\]

When \( \{ \Lambda^k \mid k = 1, 2, \ldots \} \) satisfies the condition (5), the sequence is said to be normal. It is known [10] that the conical algorithm converges to a globally optimal solution if every sequence of nested simplices is normal. In other words, to prove the convergence of the algorithm, we need only show that the norm of the cost vector is bounded from above for every \((P_k)\). This can be done at least for a problem equivalent to \((P_k)\).

By substituting \( x = Q_k \lambda \), problem \((P_k)\) can be rewritten as

\[
\begin{align*}
\text{maximize} & \quad e\lambda \\
\text{subject to} & \quad AQ_k \lambda \leq b, \quad \lambda \geq 0.
\end{align*}
\]

The dual problem is given as follows:

\[
\begin{align*}
\text{minimize} & \quad \mu b \\
\text{subject to} & \quad \mu AQ_k \geq e, \quad \mu \geq 0.
\end{align*}
\]
Let $\lambda^k$ and $\mu^k$ be optimal solutions of (6) and (7), respectively. Then $\omega^k = Q_k \lambda^k$, and by the assumption (4) we have

$$\zeta^k = e \lambda^k = \mu^k b > 1.$$ 

For the dual optimal solution $\mu^k$, let us define

$$\omega^k = Q_k \lambda^k,$$

and by the assumption (4) we have

$$\zeta^k = e \lambda^k = \mu^k b > 1.$$ 

For the dual optimal solution $\mu^k$, let us define

$$(P_k') \quad \text{maximize} \quad \mu^k A x$$
subject to

$$Ax \leq b, \quad Q_k^{-1} x \geq 0.$$ 

This problem is equivalent to $(P_k)$ in the following sense:

**Lemma 2.** An optimal solution of $(P_k')$ is $\omega^k$, and the optimal value is equal to $\zeta^k$. Conversely, if $x'$ is an optimal solution of $(P_k')$, then $x'$ is an optimal solution of $(P_k)$. \hfill \square

**Proof.** Problem $(P_k')$ is equivalent to

$$\begin{align*}
\text{maximize} & \quad \mu^k A Q_k \lambda \\
\text{subject to} & \quad A Q_k \lambda \leq b, \quad \lambda \geq 0,
\end{align*}$$

the dual of which is

$$\begin{align*}
\text{minimize} & \quad \mu b \\
\text{subject to} & \quad \mu A Q_k \geq \mu^k A Q_k, \quad \mu \geq 0.
\end{align*}$$

It is obvious that $\lambda^k$ and $\mu^k$ are feasible for (8) and (9), respectively. By the complementary slackness between 6 and 7, we have

$$\mu^k (b - A Q_k \lambda^k) = 0,$$

which reduces to the duality $\mu^k A Q_k \lambda^k = \mu^k b$ between (8) and (9). Similarly, the converse can also be proven. \hfill \square

Now let us introduce a new notion, *pseudo-nondegeneracy*, for the sequence $\{\Lambda_k \mid k = 1, 2, \ldots\}$.

**Definition 2.** The sequence of nested cones $\{\Lambda_k \mid k = 1, 2, \ldots\}$ is said to be *pseudo-nondegenerate* if there exists a subsequence $\{k_t \mid t = 1, \ldots\}$ and constant $M$ such that

$$\|\mu^k A\| \leq M, \quad k = 1, 2, \ldots.$$ 

(10)

Also, the subdivision process is *pseudo-nondegenerate* if every sequence of nested cones is pseudo-nondegenerate. \hfill \square

Let

$$\Lambda_k^+ = \{x \in \mathbb{R}^n \mid x = \sum_{j \in J_k} q_j^k \lambda_j, \lambda \geq 0\}, \quad J_k = \{j \mid \lambda_j^k > 0\}.$$ 

We can show that $\{\Lambda_k \mid k = 1, 2, \ldots\}$ is pseudo-nondegenerate even when $\Lambda_{k+1}$ is generated by subdividing $\Lambda_k$ via any $x^k \in \Lambda_k^+$ for $k = 1, 2, \ldots$. Let us refer to such a subdivision strategy as *generalized $\omega$-subdivision*. To prove the pseudo-nondegeneracy of generalized $\omega$-subdivision, we need further two lemmas, which are derived from the complementary slackness between problems (6) and (7).
Lemma 3. It holds that
\[ \mu^k Ax \geq eQ_k^{-1} x, \quad \forall x \in \Lambda_k. \]
In particular,
\[ x \in \Lambda_k^+ \Rightarrow \mu^k Ax = eQ_k^{-1} x. \]

Lemma 4. The optimal value \( \zeta^k \) of \( (P_k) \) is nonincreasing in \( k \), i.e.,
\[ \zeta^1 \geq \cdots \geq \zeta^k \geq \zeta^{k+1} \geq \cdots > 1. \]

Theorem 5. Any generalized \( \omega \)-subdivision process is pseudo-nondegenerate.

Proof. Assume that \( \| \mu^k A \| > 0 \), since otherwise there is nothing to prove, and define a halfspace
\[ H = \{ x \in \mathbb{R}^n \mid \mu^k Ax \leq \zeta^k \}. \]
If \( x \in D \), then
\[ \mu^k Ax \leq \mu^k b = \zeta^k. \]
Hence, \( D \) is a subset of \( H \). This also implies that the distance from \( 0 \), which is an interior point of \( D \) by assumption (b), to the boundary hyperplane of \( H \) is bounded from below by the distance from \( 0 \) to the boundary of \( D \), i.e.,
\[ \rho(0, \partial H) \geq \rho(0, \partial D) > 0. \]

It follows from this observation that
\[ \| \mu^k A \| \leq \zeta^k / \rho(0, \partial D), \]
Furthermore, since \( \zeta^k \) is nonincreasing in \( k \), we have
\[ \| \mu^k A \| \leq \zeta^1 / \rho(0, \partial D), \]
the right-hand-side of which is bounded from above by a constant for each instance of (1). Therefore, if we choose it as \( M \), then (10) is fulfilled for any \( \{ \Lambda_k \mid k = 1, 2, \ldots \} \) as long as \( \Lambda_{k+1} \) is generated by subdividing \( \Lambda_k \) via an \( x^k \in \Lambda_k^+ \).

4 Convergence of the subdivision process

We should remark that pseudo-nondegeneracy is a weaker condition than nondegeneracy, because the latter implies the former. Nevertheless, the normality (5) in Proposition 2 holds if the sequence \( \{ \Lambda_k \mid k = 1, 2, \ldots \} \) just satisfies the pseudo-nondegeneracy.

Let
\[ G_k = \{ x \in \mathbb{R}^n \mid eQ_k^{-1} x = 1 \} \]
\[ H_k = \{ x \in \mathbb{R}^n \mid \mu^k Ax = 1 \}, \quad H_k^+ = \{ x \in \mathbb{R}^n \mid \mu^k Ax \geq 1 \}. \]
Lemma 6. Let \(x^k\) be any point in \(\Lambda_k^+\) and \(v^k\) the intersection of the ray emanating from 0 to \(x^k\) with \(G_k\). Then
\[
v^k \in H_r^+, \quad r = 1, \ldots, k - 1.
\]
\[\square\]

Lemma 7. Let \(\{\Lambda_k \mid k = 1, 2, \ldots\}\) be a sequence of nested cones such that \(\Lambda_{k+1}\) is obtained by subdividing \(\Lambda_k\) via an \(x^k \in \Lambda_k^+\) for \(k = 1, 2, \ldots\). There exists a subsequence \(\{k_t \mid t = 1, 2, \ldots\}\) such that
\[
\{x^{k_{2s-1}}, x^{k_{2s}}\} \subset \Lambda_{k_{2s}}^+, \quad s = 1, 2, \ldots.
\] (11)
\[\square\]

Theorem 8. Let \(\{\Lambda_k \mid k = 1, 2, \ldots\}\) be a sequence of nested cones such that \(\Lambda_{k+1}\) is obtained by subdividing \(\Lambda_k\) via an \(x^k \in \Lambda_k^+\) for \(k = 1, 2, \ldots\). Then
\[
\lim_{k \to +\infty} \inf_{t \to \infty} \|\text{ext}(x^k) - v^k\| = 0,
\] (12)
where \(v^k\) is the intersection of the ray emanating from 0 to \(x^k\) with \(G_k\).

Proof. Let \(\{k_t \mid t = 1, 2, \ldots\}\) be a subsequence satisfying (11), and abbreviate \(k_t\) to \(t\). As seen in Lemma 6, while \(v^t\) is not a point of \(H_{t+1}^+\), we have
\[
v^t \in \bigcap_{r=1}^t H_r^+.
\]
Hence, according to the bounded convergence principle (see e.g., Lemma III.2 in [2], as \(t \to +\infty\), we have \(\rho(v^t, H_{t+1}^+) \to 0\), and
\[
\rho(v^t, H_{t+1}) \to 0.
\]
However, as is shown in Figure 1, we have
\[
\|\text{ext}(x^t) - v^t\| = \frac{\rho(v^t, G_{t+1})}{\rho(0, G_{t+1})} \|\text{ext}(x^t)\|
\] (13)
Suppose \(t\) is an odd number. Then both \(v^t\) and \(\text{ext}(x^t)\) are points of \(\Lambda_t^+ \cap \Lambda_{t+1}^+\) because \(x^t\) belongs to the two cones. From Lemma 3, we see that
\[
\mu^t Av^t = eQ_t^{-1}v^t, \quad \mu^t A\text{ext}(x^t) = eQ_t^{-1}\text{ext}(x^t).
\]
Therefore, if \(t = 2s - 1\), we may replace (13) by
\[
\|\text{ext}(x^{2s-1}) - v^{2s-1}\| = \frac{\rho(v^{2s-1}, H_{2s})}{\rho(0, H_{2s})} \|\text{ext}(x^{2s-1})\|.
\]
Since \(\Lambda_k\)'s are generated through a generalized \(\omega\)-subdivision process, there exists some \(M\) such that \(1/\rho(0, H_{2s}) = \|\mu_{2s}A\| < M\). Also \(\|\text{ext}(x^{2s-1})\|\) is bounded because \(\text{ext}(x^{2s-1}) \in \partial C_\gamma\), and besides \(\rho(v^{2s-1}, H_{2s}) \to 0\) as \(s \to +\infty\). Consequently, we have \(\|\text{ext}(x^{2s-1}) - v^{2s-1}\| \to 0\) as \(s \to +\infty\). 
\[\square\]
The normality of the sequence \( \{ \Lambda_k \mid k = 1, 2, \ldots \} \) generated by the usual \( \omega \)-subdivision is guaranteed as a straightforward corollary of Theorem 8, whether it is nondegenerate or not.

**Corollary 9.** Let \( \{ \Lambda_k \mid k = 1, 2, \ldots \} \) be a sequence of nested cones such that \( \Lambda_{k+1} \) is obtained by subdividing \( \Lambda_k \) via \( \omega^k \) for \( k = 1, 2, \ldots \). Then

\[
\lim_{k \to +\infty} \inf_{\infty} \| \text{ext}(\omega^k) - y^k \| = 0.
\]

\[\square\]

5 Concluding remark

In this paper, we have introduced a new concept of nondegeneracy, named pseudo-nondegeneracy, for a sequence of nested cones generated in the conical algorithm. We have shown that the \( \omega \)-subdivision process is pseudo-nondegenerate and therefore normal, even though it is still an open question whether or not the process is nondegenerate in the original sense. We have also shown in Theorem 8 that a class of generalized \( \omega \)-subdivision processes satisfies a condition similar to the normality. The usual \( \omega \)-subdivision belongs to the class and its normality is just a corollary of this theorem. However, this condition does not always guarantee the convergence of the conical algorithm unlike the normality. To make it convergent, we need a further procedure which determines a subdivision point for each cone generated in the algorithm. We will discuss the procedure in detail, in the forthcoming paper.

References


