INTERIOR-EXTERIOR POLYNOMIAL PATH-FOLLOWING APPROACH FOR CONVEX QUADRATIC PROGRAMMING

Maawiya Ould Sidi
Department of Mathematics
Faculty of Sciences
Aljouf University
Sakakah, KINGDOM OF SAUDI ARABIA

Abstract: In this paper, we consider the programming problem

\[(P): \min_{x \in \mathbb{R}^n} \{c^T x : Ax = b, \ x \geq 0\},\]

for which we construct a polynomial interior-exterior algorithm. In order to do this, we consider the sequence of penalized subproblems

\[(P_\mu): \min_{x > 0} \{f_\mu(x) = c^T x - \mu \sum_{j=1}^n \ln x_j + \frac{1}{2\mu} \sum_{i=1}^m (a_i^T x - b_i)^2\},\]

where \(x_j\) denote the \(j\)th component of \(x\) and \(a_i\) is the \(i\)th row the matrix \(A\). Problems \((P_\mu)\) are strictly convex and under some standard assumptions, have an optimal solution for all \(\mu > 0\).

AMS Subject Classification: 90C05, 90C25, 90C46, 90C51
Key Words: interior-exterior algorithm, primal-dual problem, convex-quadratic programming

1. Introduction

Since the publication of the Karmarkar’s polynomial time algorithm in [8, 9], a wide variety of interior point methods has been presented to solve efficiently linear programming problems in polynomial time. A number of individuals have
contributed to the field in many different ways, including theoretical development, computational aspects, and exploration of new algorithms. An excellent survey appears in the work of Gonzaga [7]. Many others methods are focused on interior point methods [1, 2, 5, 6, 10, 12]. In majority, interior point methods are based on the application of the logarithmic barrier function technique. The logarithmic barrier function method consists of examining the family of problems

\[
\min_{x>0} \{c^T x - \mu \sum_{j=1}^{n} \ln x_j : Ax = b\},
\]

where \(\mu > 0\) is the barrier penalty parameter.

In this paper, we construct an interior-exterior algorithm by considering the family of penalized subproblems

\[
(P_\mu) : \min_{x>0} \{f_\mu(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j + \frac{1}{2\mu} \sum_{i=1}^{m} (a_i.x - b_i)^2\},
\]

where \(x_j\) is the \(j\)th component of \(x\) and \(a_i\) is the \(i\)th row of the matrix \(A\).

Problems \((P_\mu)\) are strictly convex and, under some standard assumptions, have an optimal solution for all \(\mu > 0\). On other hand the solution \(\omega(\mu) = (x(\mu), y(\mu), z(\mu))\) of the system corresponding to the Karush-Kuhn-Tucker (K-K-T) conditions gives a dual feasible solution and a primal \(\mu\)-feasible solution. Finally, every accumulation point \(\omega^*\) of the sequence \(\omega(\mu)\) is such that \(x^*\) is an optimal solution of the primal problem

\[
(P) : \min \{c^T x : Ax = b, x \geq 0\}
\]

and \((y^*, z^*)\) is an optimal solution of the dual

\[
(D) : \max \{b^T y : A^T y = z = c, z \geq 0\},
\]

the constructed algorithm is polynomial and its complexity is \(O(n^{3.5}L)\). In the next section, we present a theoretical analysis of our approach. We describe the properties of the penalized subproblems, we define a new notion of pseudo-gap in order to characterize the solutions generated with iterations, and we present a theoretical algorithm of resolution.

2. Main Result

We consider the pair of the standard form linear program and its dual

\[
(P) : \min_{x \in \mathbb{R}^n} \{c^T x\}
\]

\[
(D) : \max \{b^T y : A^T y = z = c, z \geq 0\},
\]
\[
(D) : \max_{y \in \mathbb{R}^m, x \in \mathbb{R}^n} \{b^T y : A^T y + z = c, \ z \geq 0\},
\]
and we impose the following assumptions:

(H1) The set \( S \equiv \{x \in \mathbb{R}^n : Ax = b, \ x > 0\} \) is non-empty.

(H2) The set \( T \equiv \{(y, z) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + z = c, \ z > 0\} \) is non-empty.

(H3) \( A \) is an \( n \times m \) matrix of full rank and \( m \leq n \), i.e. \( \text{rang}(A) = m \).

(H4) The dual feasible region is bounded (Freund [3]): Let \( \alpha \) the smallest positive number such that \( \|y\|^2 \leq \alpha \) for all \( y \in \mathbb{R}^m \) satisfying \( A^T y \leq c \).

We say that points in the sets \( S \) and \( T \) are interior feasible solutions of problems \( P \) and \( D \) respectively.

### 2.1. Penalized Subproblems: Definition and Properties

To solve the primal problem \( P \), we consider the following penalized problem:

\[
(P_\mu) : \min_{x > 0} \{f_{\mu}(x) = c^T x - \mu \sum_{j=1}^n \ln x_j + \frac{1}{2\mu} \sum_{i=1}^m (a_i x - b_i)^2\}, \text{ where } \mu > 0.
\]

Observe that \( f_\mu \), the objective function of problem \( P_\mu \), is a strictly convex function.

Indeed, it suffices to remark that the function \( f_1 \) defined by:

\[
f_1 : x \in \mathbb{R}^n \rightarrow c^T x - \frac{1}{2\mu} \sum_{i=1}^m (a_i x - b_i)^2
\]

is convex and that the function \( f_2 \) given by:

\[
f_2 : x \in \mathbb{R}^n \rightarrow -\mu \sum_{j=1}^n \ln x_j
\]

is strictly convex for all \( \mu > 0 \). Accordingly \( f_\mu \) which is the function \( f_1 + f_2 \) is strictly convex. This implies that problems \( P_\mu \) has at most one minimum, and that this minimum, if it exists, is global and completely characterized by the K-K-T optimality conditions:

\[
c - \mu X^{-1} e + \frac{1}{\mu} A^T (A x - b) = 0 \quad \text{and} \quad x > 0,
\]
where \( e \) denotes the \( \mathbb{R}^n \) unit vector i.e. \( e = (1, \ldots, 1)^T \).

By introducing two vectors \( y \in \mathbb{R}^m \) et \( z \in \mathbb{R}^n \), this system can be rewritten in an equivalent way as:

\[
(S_\mu) \begin{cases}
ZXe - \mu e = 0 \\
Ax + \mu y - b = 0, \ x > 0 \\
A^T y + z - c = 0
\end{cases}
\]

We denote the unique solution of \((S_\mu)\) by \( \omega(\mu) = (x(\mu), y(\mu), z(\mu)) \).

A necessary and sufficient condition for problem \((P_\mu)\) to have an optimal solution for all \( \mu > 0 \) is given by the following propositions.

**Proposition 1.** Let \( \mu > 0 \) be given and assume Assumption \((H_1)\) holds. Then problem \((P_\mu)\) has an optimal solution if and only if the set of optimal solutions of problem \((P)\) is non-empty and bounded.

**Proof.** \((\Rightarrow)\) If problem \((P_\mu)\) has an optimal solution \( x(\mu) \) then the K.K.T optimality conditions for \((P_\mu)\) hold and problem \((D)\) is feasible. Therefore problems \((P)\) and \((D)\) have optimal solutions and the optimal values \( v(P) \) and \( v(D) \) of \((P)\) and \((D)\) respectively are equal: \( v(P) = v(D) \). Let \( x^* \) an optimal solution of \((P)\).

Denote \( Opt(P) = \{ x \in \mathbb{R}^n : Ax = b, \ x \geq 0 \text{ and } c^T x = v(P) \} \).

Assume that \( Opt(P) \) is not bounded. Then there is a non zero vector \( d \geq 0 \) in \( N(A) \), such that: \( u(\lambda) = x^* + \lambda d \in Opt(P), \ \forall \lambda > 0 \), where \( N(A) \) denote the null space of the matrix \( A \).

Now, let \( h = x(\mu) + \lambda d \), one has: \( Ah = Ax(\mu) + \lambda A d = Ax(\mu) \) and \( h > 0 \) since \( x(\mu) > 0 \), \( \lambda d \geq 0 \) and \( Ad = 0 \).

Hence one deduces that:

\[
c^T x(\mu) - \mu \sum_{j=1}^{n} \ln x_j(\mu) + \frac{1}{2\mu} \sum_{i=1}^{m} (a_i x(\mu) - b_i)^2 > c^T h - \mu \sum_{j=1}^{n} \ln h_j + \frac{1}{2\mu} \sum_{i=1}^{m} (a_i h - b_i)^2.
\]

Indeed

- \( c^T x(\mu) = c^T h \) since \( c^T d = 0 \)
- \( a_i x(\mu) = a_i h = a_i x(\mu) + \lambda a_i d \) since \( d \in N(A) \) \( \Rightarrow a_i d = 0, \ \forall i \)
• \( x_j(\mu) \leq x_j(\mu) + \lambda d_j \) for all \( j \in \{1, \ldots, n\} \) and \( x_k(\mu) < x_k(\mu) + \lambda d_k \) for at least one \( k \in \{1, \ldots, n\} \) since \( d \neq 0 \).

Which contradicts the fact that \( x(\mu) \) is optimal. So \( \text{Opt}(P) \) is bounded.

(\( \Leftarrow \)) One can easily verify that the set of optimal solutions of problem \( (P) \) is non-empty and bounded if and only if \( \{d \in \mathbb{R}^n : c^T d \leq 0, \ A d = 0, \ d \geq 0\} = \{0\} \).

Let \( x > 0 \), then

\[
f_{\mu}(x) = \|x\| \frac{c^T x}{\|x\|} - \mu \sum_{j=1}^{n} \ln \left( \frac{x_j}{\|x\|} \right) - n \mu \ln \|x\| + \frac{\|x\|^2}{2\mu} \left\| \frac{Ax}{\|x\|} - \frac{b}{\|x\|} \right\|^2
\]

If \( \|x\| \to +\infty \), then \( \frac{x}{\|x\|} \to \overline{x} \neq 0, \ \|\overline{x}\| \geq 0 \).

If we assume that \( f_{\mu} \) the objective function of problem \( (P_{\mu}) \) goes to \( -\infty \) then we must have \( \frac{c^T x}{\|x\|} \to c^T \overline{x} \leq 0 \) and

\[
\left\| \frac{Ax}{\|x\|} - \frac{b}{\|x\|} \right\|^2 \to \|A\overline{x}\|^2 = 0
\]

with \( \overline{x} \neq 0 \). Contradiction.

\[\Box\]

**Corollary 2.** If problem \( (P_{\mu}) \) has a solution for some \( \mu > 0 \) then it has a solution for all \( \mu > 0 \).

The role played by Assumption \((H_2)\) is now provided by the following result.

**Proposition 3.** Assume that problem \((P)\) is feasible. Then the set of optimal solutions of problem \((P)\) is non-empty and bounded if and only if Assumption \((H_2)\) holds, that is, the set of interior feasible solutions of the dual problem \((D)\) is non-empty.

The proof of 3 results from the duality theory for linear programming.

As a consequence of the two previous propositions, we have the following corollaries.
Corollary 4. Under Assumptions \((H_1)\) and \((H_2)\), problem \((P_{\mu})\) (and consequently system \((S_{\mu})\)) has a unique solution \(x(\mu) = (\omega(\mu), y(\mu), z(\mu))\) for all \(\mu > 0\).

Corollary 5. Let \(\mu > 0\) fixed in system \((S_{\mu})\). Then the point \((y(\mu), z(\mu))\) is an interior feasible solution for the dual problem \((D)\), that is, \((y(\mu), z(\mu)) \in T\).

We conclude this section by mentioning a classical result due to Fiacco and Mc Comick providing that every accumulation point of the sequences \(x(\mu)\) and \((y(\mu), z(\mu))\) is an optimal solution of problems \((P)\) and \((D)\) respectively.

2.2. Pseudo-Gap: Definition and Properties

Since the point \(x(\mu)\) is not a feasible solution of the primal problem \((P)\), the classical definition of the duality does not allow us to characterize in a satisfactory way the solutions \(x(\mu), y(\mu), z(\mu)\) generated with iterations. This leads us to introduce the following definition.

Definition 6. The pseudo-gap at point \(\omega(\mu)\) is the quantity given by:

\[
\triangle(\mu) = |c^T x(\mu) - b^T y(\mu)| + \frac{1}{\mu} \|Ax(\mu) - b\|^2.
\]

Proposition 7. One has: \(\triangle(\mu) = |x(\mu)^T z(\mu) - \mu\|y(\mu)\|^2| + \mu\|y(\mu)\|^2\)

moreover

\[
\triangle(\mu) = \begin{cases} 
\mu n & \text{if } \|y(\mu)\|^2 \leq n, \\
\mu(2\|y(\mu)\|^2 - n) & \text{if } \|y(\mu)\|^2 > n.
\end{cases}
\]

Proof. Using the two last equations in system \((S_{\mu})\), one can easily verify the first above relation whereas the second expression is deduced from the first equation in system \((S_{\mu})\). \(\square\)

Proposition 8. Let \(\triangle(\mu_k)\) denote the pseudo-gap at point \(\omega(\mu_k) = (x^k, y^k, z^k)\), which is solution of system \((S_{\mu_k})\), and \(\beta = \sup(2\alpha, 2n) \) [ \(\alpha\) is defined by Assumption \((H_4)\)] . Then one has:

\(\triangle(\mu_k) \leq (\beta - n)\mu_k\).

Proof. If \(\triangle(\mu_k) = n\mu_k\) then \(\triangle(\mu_k) = (2n - n)\mu_k \leq (\beta - n)\mu_k\).

Else \(\triangle(\mu_k) = (2\|y^k\|^2 - n)\mu_k \leq (2\alpha - n) \leq (\beta - n)\mu_k\). \(\square\)
Remark 9. Since $(\beta - n)$ is a positive constant and $(\mu_k)_{k \in \mathbb{N}}$ is a positive sequence strictly decreasing to zero, the pseudo-gap convergence as well to zero as $\mu_k$ goes to zero.

On the other hand, it is clear that $\Delta(\mu)$ does not converge to zero unless $x(\mu)$ and $(y(\mu), z(\mu))$ converge to an optimal solutions of $(P)$ and $(D)$ respectively.

2.3. Theoretical Algorithm

<table>
<thead>
<tr>
<th>Initialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Let $x^0 &gt; 0$ and $\mu_0 &gt; 0$,</td>
</tr>
<tr>
<td>• Let $\epsilon &gt; 0$ be a tolerance for the pseudo-gab,</td>
</tr>
<tr>
<td>• Set $k := 0$.</td>
</tr>
</tbody>
</table>

| While $\Delta(\mu_k) > \epsilon$ Do |
| • Choose $\mu_{k+1} < \mu_k$, |
| • Determiner $(x(\mu), y(\mu), z(\mu))$ solution of $(S_\mu)$, |
| • Set $k := k + 1$. |

References


