Complexity of interior-point algorithms: tropical compuations

Mohamed Amine Bennouna, Yassine El Maazouz

December 2018

Our deep gratitude goes to prof. ***Xavier Allamigeon*** *& prof.* ***Stéphane Gaubert*** *for introducing us to the wonderful world of tropical geomerty and for all their time and dedication to help us complete this work.*

Abstract

Interior-point methods are a very import class of algorithms that solve linear and semi-definite optimization problems. Very successful in practice, they are widely believed to be strongly polynomial. In this report, we follow on a recently published paper proving that log-barrier interior point methods are not strongly polynomial by presenting a counter-example. We analyse the behaviour of a predictor-corrector algorithm on the counter-example instances to confirm the theoretical results. In particular, we present an explicit method of constructing an initial point for the algorithm over the field of generalized Puiseux series based on tropical characteristics of the central path.

1 Introduction and context

Let n , $m \in \mathbb{N}$ be two positive integers, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ two vectors and $A \in \mathbb{R}^{m\times n}$ a real $m \times n$ matrix. We are interested in solving linear programs which are optimization problems of the following form:

$$\text{minimize} \quad \langle c, x\rangle \quad \text{subject to } Ax \leq b \quad \text{and} \quad x \geq 0$$

where c is the cost vector and $x \geq 0$ mean that $x_i \geq 0$ for $1 \leq i \leq n$ (x is in the positive orthant).
By adding slack variables $w \in \mathbb{R}^m_{\geq 0}$ to x the linear program (P) can be transformed to the following form:

$$(P) \qquad \inf_{x\in\mathbb{R}^n, w\in\mathbb{R}^m} \langle c, x\rangle \quad \text{subject to } Ax + w = b \quad \text{and} \quad (x, w) \geq 0$$

We denote by $\mathcal{F}_P$ feasible set of the problem (P) is the set of vectors x, w that satisfy the constraints, formally:

$$\mathcal{F}_p = \{(x, w) \in \mathbb{R}^n \times \mathbb{R}^m : Ax + w = b, x \geq 0, w \geq 0\}$$

We call the strictly feasible set $\mathcal{F}^s_P$ the subset of $\mathcal{F}_P$ that corresponds to strict positivity constraints over x and w:

$$\mathcal{F}^s_p = \{(x, w) \in \mathbb{R}^n \times \mathbb{R}^m : Ax + w = b, x > 0, w > 0\}$$

.

The dual program of (P) is denoted (D) and is written as follows:

$$(D) \quad \sup_{y\in\mathbb{R}^m, s\in\mathbb{R}^n} \langle b, y\rangle \quad \text{subject to } A^T y + s = c \quad \text{and} \quad (y,s) \geq 0$$

This dual program also gives rise to the corresponding feasible and strictly feasible sets :

$$\mathcal{F}_D = \{(y,s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c,\ y \geq 0,\ s \geq 0\}$$
$$\mathcal{F}_D^s = \{(y,s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c,\ y > 0\ s > 0\}$$

We define the primal-dual feasible set $\mathcal{F}$ and strictly feasible set $\mathcal{F}^s$ as follows:

$$\mathcal{F} = \mathcal{F}_P \times \mathcal{F}_D \text{ and } \mathcal{F}^s = \mathcal{F}_P^s \times \mathcal{F}_D^s$$

A familiar result in linear programming (see [2] for more details) states that $(x,w) \in \mathbb{R}^n \times \mathbb{R}^m$ is solution to (P) if and only if there exist a couple $(y,s) \in \mathbb{R}^m \times \mathbb{R}^n$ such that :

$$\begin{cases} A^T y + s = c,\ s \geq 0,\ y \geq 0 \\ Ax + w = b,\ x \geq 0\ w \geq 0 \\ \langle x, s\rangle = 0 \end{cases} \tag{1}$$

The same result holds for the dual: (y,s) is an optimal solution of (D) if and only if there exist $(x,w) \in \mathbb{R}^n \times R^m$ such that (1) holds. In this case we denote by $z = (x,w,y,s)$ the primal-dual solution of (1).

1.1 Logarithmic barrier and interior point algorithms

Positivity constraints are the main difficulty of the Primal-Dual linear program summarized in (1). To get rid of these conditions we penalize the points on the borders of the positive orthant by adding a barrier to the cost function that pushes the optimal solutions away from the boundaries of the feasible set $\mathcal{F}$ into the set $\mathcal{F}_s$. Formally we define a family of programs parametrized by $\mu > 0$ as follows:

$$(P_\mu) \begin{cases} \inf\ \langle c, x\rangle + \mu.\mathrm{lb}(x) \\ Ax + w = b, x \geq 0, w \geq 0 \end{cases} \qquad \text{where: } \mathrm{lb}(x) = \begin{cases} -\sum_{i=1}^n \log(x_i), & \text{if } x > 0 \\ +\infty, & \text{else} \end{cases}$$

The *log-barrier* function lb, *pushes* the solution into the positive orthan by *penalizing* the objective function when we get close to the borders, as $\mathrm{lb}(x) \to +\infty$ when $x_i \to 0$.

The dual of this problem is written as:

$$(D_\mu) \begin{cases} \sup\ \langle b, y\rangle - \mu \mathrm{lb}(s) \\ A^T y + s = c, y \geq 0, s \geq 0 \end{cases}$$

The primal-dual problem can be summarized (as we did above) in the following problem:

$$\begin{cases} A^T y + s = c,\ s > 0 \\ Ax + w = b,\ x > 0, w > 0 \\ x_i s_i = \mu, \forall 1 \leq i \leq n \end{cases} \tag{2}$$

Solving $(P_\mu) - (D_\mu)$ for all $\mu > 0$ gives an application:

$$C : \mu \in]0, +\infty[\to z(\mu) := (x(\mu), w(\mu), s(\mu), y(\mu)) \in \mathcal{F}^s$$

We recall that $\mathcal{F}^s := \mathcal{F}_P^s \times \mathcal{F}_D^s$ is the primal-dual strictly feasible set.

Notice that (P_μ) and (D_μ) becomes (P) and (D) when $\mu \to 0$ and that $(C(\mu))_\mu$ is an algebraic curve. We show, under reasonable hypothesis, that we have as we may expect $z(\mu) \underset{\mu \to 0}{\longrightarrow} z^* = (x^*, w^*, y^*, s^*)$ where z^* is optimal primal-dual solution of the original problem. Therefore, the algebraic curve $(C(\mu))_{\mu \geq 0}$ (defined by the polynomial equations in (4)) is is path converging to the optimal solution of the problem as μ tends to 0 which is called the ***central path***. Log-barrier interior point algorithms consist of following this path to the optimal solution.

We now present as special type of interior point algorithm that we will will be using throughout in all that follows.

1.2 Predictor-Corrector algorithm

In this section we present the predictor corrector algorithm. We should stress that we consider the following linear program with equality constraints:

$$\inf_{x \in \mathbb{R}^n} \quad \langle c, x \rangle \quad \text{subject to } Ax = b \quad \text{and} \quad x \geq 0$$

whose dual problem is of the form:

$$\sup_{y \in \mathbb{R}^m, s \in \mathbb{R}^n} \quad \langle b, y \rangle \quad \text{subject to } A^T y + s = c \quad \text{and} \quad s \geq 0$$

We can transform a linear problem with inequality constraints to a problem of the above form by adding slack variables. We now describe the predictor corrector algorithm over a problem with equality constraints

Predictor-Corrector algorithm is an example of interior point methods. It is based on a opportunisticway of following the central path towards the optimal solution. On every step of the algorithm, we start by following a Newton direction to the optimal solution by making sure that we don't deviate too much from the central path (prediction) and then correct the path by getting closer to the central path (correction) which yields a piecewise approximation of the central path.

Formally, let's start by defining the *duality measure* that quantify how close we are from the optimal solution. For a given $z \in \mathcal{F}^s$, we define the duality measure as:

$$\bar{\mu}(z) = \frac{\langle x, s \rangle}{n}$$

Notice that $z \in \mathcal{F}^s$ is in the central path, if and only if $x * s = \bar{\mu}(z)e$, where $*$ is the Hadamard product of vectors ($(x * s)_i = x_i s_i$), and e is the vector in $\mathbb{R}^n$ whose entries are all 1.

In order to prevent the prediction step so as not to go so far from the central path, we need to define a neighbourhood around the central path. We will be using two neighbourhoods: one to limit the prediction step, and one to define a tolerance for the correction step. The neighbourhoods are defined by a parameter θ and can be of two types:

The norm-2 neighbourhood:

$$V_2(\theta) := \{z \in \mathcal{F}^s : \|x * s - \bar{\mu}(z)e\|_2 \leq \theta\bar{\mu}(z)\}$$

The norm-∞ neighbourhood:

$$V_\infty^-(\theta) := \{z \in \mathcal{F}^s : \|x * s - \bar{\mu}(z)e\|_\infty \leq \theta\bar{\mu}(z)\}$$

where $\|v\|_\infty = \max(0, \max_i(-v_i))$.

The norm-∞ neighbourhood can also be rewritten as:

$$V_\infty^-(\theta) := \{z \in \mathcal{F}^s : x * s \geq (1-\theta)\bar{\mu}(z)e\}$$

The following algorithm works for both types of neighbourhood and we refer to the neighbourhood as $V(\theta)$.

Algorithm 1 Predictor-corrector algorithm

We consider two sets $V(\theta)$ and $V(\theta')$ with $0 < \theta' < \theta$.
Suppose that the actual point is $z := (x, w, s, y) \in V(\theta') \subset \mathcal{F}^s$. The following point $z_+ \in V(\theta')$ is obtained by two steps.

1. **Prediction**
   (a) Compute Newton direction d from z toward the optimal solution such that we stay in the feasible set $\mathcal{F}^s$.
   (b) Compute the largest jump $\alpha \in [0, 1]$ such that $z' = z + \alpha d \in V(\theta)$.
   (c) Prediction point: z'.
2. **Correction**
   (a) Compute Newton direction d' from z' toward the central path such that we stay in the feasible set $\mathcal{F}^s$ and get closer to the narrow neighbourhood $V(\theta)$.
   (b) Correction point: $z_+ = z' + d'$.

In this part we explain how the Newton directions are computed.

1.a We compute $d = \begin{pmatrix} dx \\ dy \\ ds \end{pmatrix}$ as a solution of [2]:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ ds \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -Xs \end{pmatrix}$$

Step 1.b

When $V = V_2$, we compute α by solving: $\|x(\alpha) * s(\alpha) - \bar{\mu}(z(\alpha))e\|_2^2 = \theta'^2\bar{\mu}(z(\alpha))^2$, where $z(\alpha) = z + \alpha d$.

When $V = V_\infty^-$, we need to compute:

$$\begin{aligned} \alpha^* &= \max\{\alpha \in [0, 1] \mid X(\alpha)s(\alpha) \geq (1-\theta)\bar{\mu}(z(\alpha))e\} \\ &= \max\{\alpha \in [0, 1] \mid \min_i x_i(\alpha)s_i(\alpha) = (1-\theta')\bar{\mu}(z(\alpha))\} \end{aligned} \tag{3}$$

In order to do that, we compute $\alpha^{**} = \min_i \min\{\alpha_i \in [0,1] \mid x_i(\alpha_i)s_i(\alpha_i) = (1-\theta')\bar{\mu}(z(\alpha_i))\}$ which is easier to compute (solving n second degree polynomial equations). Given lemma 2 in [1], we have $z + \alpha d \in V$ for all $\alpha \in [0, \alpha^*]$ and therefore $\alpha^* = \alpha^{**}$.

Step 2.b We compute $d' = \begin{pmatrix} d'x \\ d'y \\ d's \end{pmatrix}$ as a solution of [2]:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} d'x \\ d'y \\ d's \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \bar{\mu}(z)e - Xs \end{pmatrix}$$

We use the values $\theta' = \frac{1}{2}$ and $\theta = \frac{1}{4}$ which are values for which the algorithm in known to converge (these values are not unique). For more details on the predictor corrector algorithm, we refer the reader to of the book [2, chapter 18] by Jean-Charles Gilbert.

2 Log-barrier interior point methods are not strongly polynomial

In their article, Stéphane Gaubert, Xavier Allamigeon and their co-authors prove that log-barrier Interior-Point methods are not strongly polynomial by presenting a counter example instance inspired by a Game-Theory problem.

The counter example is the follwing:

Given an integer number $r > 1$ and an Archimedean parameter t we consider the following two parameter Linear Program which we refer to as $\mathbf{LW}_r(t)$:

$$\begin{aligned}
\text{Minimize} & x_1 \\
\text{Subject to} & x_1 \leq t^2 \\
& x_2 \leq t \\
& x_{2j+1} \leq t\, x_{2j-1}\,,\ x_{2j+1} \leq t\, x_{2j}, \qquad 1 \leq j < r \\
& x_{2j+2} \leq t^{1-1/2^j}(x_{2j-1} + x_{2j}), \qquad 1 \leq j < r \\
& x_{2r-1} \geq 0\,,\ x_{2r} \geq 0
\end{aligned} \tag{4}$$

$\mathbf{LW}_r(t)$ has $2r$ variables and $3r+1$ constraints and can be transformed to fit the form of (P) by adding slack variables $(w_j)_{1 \leq j \leq r}$. When transformed, $\mathbf{LW}_r(t)$ has $5r+1$ variables and $3r+1$ constraints.

$\mathbf{LW}_r(t)$ can be rewritten in a matricial form as follows:

$$\begin{aligned}
\text{Minimize} \quad & c.x \\
\text{Subject to } & A(t)x \leq b(t), \text{ and } x \geq 0
\end{aligned}$$

References

[1] Log-Barrier Interior Point Methods Are Not Strongly Polynomial X.Allamigeon, P.Benchimol, S.Gaubert, and M.JOSWIG

[2] https://who.rocq.inria.fr/Jean-Charles.Gilbert/ensta/optim.html