1

Complexity of the Linear Programming Feasibility Problem

1.1 Linear Programming

We start with a brief review of the basic concepts of linear programming, compare [1]. Linear programs are usually stated in the following standard form

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

where \(A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n\) are the given data and we look for an optimal vector \(x \in \mathbb{R}^n\). Here and in what follows \(x \geq 0\) means \(x_i \geq 0\) for all \(i\) and \(x > 0\) means \(x_i > 0\) for all \(i\). We call (P) feasible if there exists \(x \in \mathbb{R}^n\) such that \(Ax = b, \ x \geq 0\). A feasible problem (P) is called bounded if the minimum of the objective function is finite. Otherwise it is called unbounded.

Given the data \(A, b, c\) we consider the closely related optimization problem

\[(D) \max b^T y \text{ subject to } A^T y \leq c,\]

where \(y \in \mathbb{R}^m\). This problem is called the dual problem of (P). Here we can also speak about feasibility and boundedness.

Suppose now that (P) and (D) are both feasible, say

\[Ax = b, \ x \geq 0, \ A^T y \leq c\]

for some \(x \in \mathbb{R}^n, y \in \mathbb{R}^m\).

Introducing the vector \(s := c - A^T y\) of slack variables we have \(A^T y + s = c\) and \(s \geq 0\). Then

\[c^T x - b^T y = (s^T + y^T A)x - b^T y = s^T x + y^T (Ax - b) = s^T x \geq 0.\]

It follows that (P) and (D) are both bounded and \(\max b^T y \leq \min c^T x\). The fundamental duality theorem of linear programming states that actually equality holds. A proof can be found e.g. in [1].

**Theorem 1.1 (Duality Theorem of Linear Programming)**

(i) The problem (P) is bounded iff (D) is bounded. In this case the objective values of (P) and (D) are equal.
(ii) If (P) is unbounded then (D) is infeasible. If (D) is unbounded, then (P) is infeasible.

Often it is useful to attempt to solve (P) and (D) simultaneously. Consider the polyhedral set \( \mathcal{F} \) of feasible points \( z = (x, y, s) \in \mathbb{R}^{n+m+n} \) satisfying
\[
Ax = b, \quad A^Ty + s = c, \quad x \geq 0, \quad s \geq 0.
\]
We note that \( \mathcal{F} \) is convex. It follows from (1.1) and the Duality Theorem 1.1 that for \((x, y, s) \in \mathcal{F}\), \(x\) is an optimal solution of (P) and \(y\) is an optimal solution of (D) iff the complementary slackness condition
\[
x_i s_i = 0 \quad i = 1, 2, \ldots, n
\]
holds.

### 1.2 Primal-Dual Interior Point Methods: Basic Ideas

The most common method to solve linear programs is Dantzig’s simplex method. This method relies on the geometry of the polyhedron of solutions and follows a path of vertices on the boundary of the polyhedron. By contrast, interior point methods follow a path in the interior of the polyhedron, hence the name. The path is a nonlinear curve that is approximately followed by a variant of Newton’s method.

More specifically, **primal-dual interior point methods** search for solutions of the system
\[
Ax = b, \quad A^Ty + s = c, \quad x \geq 0, \quad s \geq 0, \quad x_1 s_1 = 0, \ldots, x_n s_n = 0
\]
by following a certain curve in the strictly feasible set \( \mathcal{F}^\circ \subseteq \mathbb{R}^{n+m+n} \) defined by
\[
Ax = b, \quad A^Ty + s = c, \quad x > 0, \quad s > 0.
\]
Note that (1.4) is only mildly nonlinear (quadratic equations \( x_is_i = 0 \)). It is the nonnegativity constraints that form the main source of difficulty. For a parameter \( \mu > 0 \) we add now to (1.5) the additional constraints
\[
x_1 s_1 = \mu, \ldots, x_n s_n = \mu.
\]
One calls \( \mu \) the duality measure. Under certain genericity assumptions, there is exactly one strictly feasible solution \( \zeta_\mu \in \mathcal{F}^\circ \) satisfying (1.6) and such that \( \zeta = (x, y, s) = \lim_{\mu \to 0} \zeta_\mu \) exists. Then it is clear that \( \zeta \in \mathcal{F} \) and \( x_is_i = 0 \) for all \( i \). Hence \( \zeta \) is a desired solution of the primal-dual optimization problem.

We postpone the proof of the next theorem to Section 1.3.

**Theorem 1.2** Suppose that \( m \leq n \), \( \text{rk} A = m \) and that there is a strictly feasible point, i.e., \( \mathcal{F}^\circ \neq \emptyset \). Then for all \( \mu > 0 \) there exists a uniquely determined point \( \zeta_\mu = (x_\mu, y_\mu, s_\mu) \in \mathcal{F}^\circ \) such that \( x_is_i = \mu \) for \( i = 1, 2, \ldots, n \).
Definition 1.3 The central path $C$ of the primal-dual optimization problem given by $A, b, c$ is the set

$$C = \{ \zeta_\mu : \mu > 0 \}.$$  

Suppose we know $\zeta_{\mu_0}$ for some $\mu_0 > 0$. The basic idea is to choose a sequence of parameters $\mu_0 > \mu_1 > \mu_2 > \ldots$ converging to zero and to successively compute approximations $z_k$ of $\zeta_k := \zeta_{\mu_k}$ for $k = 0, 1, 2, \ldots$ until a certain accuracy is reached. In most cases one chooses $\mu_k = \sigma^k \mu$ with a centering parameter $\sigma \in (0, 1)$.

It is useful to define a duality measure for any $z = (x, y, s) \in F^\circ$ by

$$\mu(z) := \frac{1}{n} \sum_{i=1}^{n} x_i s_i = \frac{1}{n} s^T x.$$  

How can we get the approximations $z_k$? This is based on Newton’s method, one of the most fundamental methods in computational mathematics. Consider the map $F: \mathbb{R}^{n+m+n} \to \mathbb{R}^{n+m+n}$,

$$z = (x, y, s) \mapsto F(z) = (A^T y + s - c, Ax - b, x_1 s_1, \ldots, x_n s_n).$$

We note that $\{ \zeta_\mu \} = F^{-1}(0, 0, \mu e)$, where $e := (1, \ldots, 1) \in \mathbb{R}^n$, by Theorem 1.2. The Jacobian matrix of $F$ at $z$ equals

$$DF(z) = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix},$$

where here and in the following we set

$$S = \text{diag}(s_1, \ldots, s_n), \ X = \text{diag}(x_1, \ldots, x_n).$$

Depending on the context, $z, e$ etc. should be interpreted as column vectors.
Lemma 1.4 If \( m \leq n \) and \( \text{rk} \, A = m \), then \( DF(z) \) is invertible, provided \( s,x_i \neq 0 \) for all \( i \).

Proof. By elementary column operations we can bring the matrix \( DF(z) \) to the form
\[
\begin{bmatrix}
D & A^T & I \\
A & 0 & 0 \\
0 & 0 & X
\end{bmatrix},
\]
where \( D = \text{diag}(-s_1 x_1^{-1}, \ldots, -s_n x_n^{-1}) \). It is therefore sufficient to show that the matrix \( [D \ A^T A \ 0] \) is invertible. Such matrices are called of Karush-Kuhn-Tucker type. Suppose that \( [D \ A^T A \ 0] \begin{bmatrix} x \\ y \end{bmatrix} = 0 \), that is, \( Dx + A^T y = 0, Ax = 0 \). It follows that
\[
0 = \begin{bmatrix} x^T \\ y^T \end{bmatrix} [D \ A^T A \ 0] \begin{bmatrix} x \\ y \end{bmatrix} = x^T Dx + (Ax)^T y = x^T Dx.
\]
As \( D \) is negative definite, it follows \( x = 0 \). Hence \( A^T y = 0 \). Therefore, \( y = 0 \), as \( \text{rk} \, A = m \).

We continue with the description of the basic algorithmic idea. Choose \( \mu_k = \sigma^k \mu_0 \) and set \( \zeta_k = \zeta_{\mu_k} \). Then \( F(\zeta_k) = (0, 0, \mu_k e) \) for all \( k \in \mathbb{N} \). A first order approximation gives
\[
F(\zeta_{k+1}) \approx F(\zeta_k) + DF(\zeta_k)(\zeta_{k+1} - \zeta_k).
\]
Suppose that \( z_k = (x, y, s) \in F^\circ \) is an approximation of \( \zeta_k \). Then \( F(z_k) = (0, 0, x_1 s_1, \ldots, x_n s_n) = (0, 0, X^T s e) \). We obtain from (1.7), replacing the unknowns \( \zeta_k \) by \( z_k \),
\[
(0, 0, \mu_{k+1} e)^T = F(\zeta_{k+1}) \approx F(z_k) + DF(z_k)(z_{k+1} - z_k).
\]
This leads to the definition
\[
z_{k+1} := z_k + DF(z_k)^{-1}(0, 0, \mu_{k+1} e - X^T s e)^T
\]
of the approximation of \( \zeta_{k+1} \). This vector is well defined due to Lemma 1.4. Put \( z_{k+1} = z_k + (\Delta x, \Delta y, \Delta s) \). Then
\[
(A^T \Delta y + \Delta s, A \Delta x, \Delta x + \Delta s) = DF(z_k)(\Delta x, \Delta y, \Delta s)^T = (0, 0, \mu_{k+1} e - X^T s e)^T,
\]
hence \( A^T \Delta y + \Delta s = 0, A \Delta x = 0 \), which implies \( A^T (y + \Delta y) + (s + \Delta s) = e, A(x + \Delta x) = b \). We have shown that \( z_{k+1} \) satisfies the equalities in (1.5). By a suitable choice of the parameter \( \sigma \) we will see that one can achieve that \( z_{k+1} \) also satisfies the strict inequalities in (1.5), that is \( z_{k+1} \in F^\circ \).

Summarizing, the framework for a primal-dual interior point method is the following:
Algorithm 1.1: Primal-Dual IPM

Input: $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ s.t. $\text{rk} A = m \leq n$.
Choose starting point $z_0 = (x^0, y^0, s^0) \in \mathcal{F}^o$ with duality measure $\mu_0$.
Choose centering parameter $\sigma \in (0, 1)$.

for $k = 0, 1, 2, \ldots$

Solve

$$
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S^k & 0 & X^k
\end{bmatrix} \begin{bmatrix}
\Delta x^k \\
\Delta y^k \\
\Delta s^k
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
\sigma^{k+1} \mu_0 e - X^k S^k e
\end{bmatrix},
$$

where $X^k = \text{diag}(x_1^k, \ldots, x_n^k)$, $S^k = \text{diag}(s_1^k, \ldots, s_n^k)$.

Set

$$(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k, y^k, s^k) + (\Delta x^k, \Delta y^k, \Delta s^k).$$

until some stopping criterion is matched.

1.3 Existence and uniqueness of central path

We provide here the proof of the fundamental Theorem 1.2, following [4].

Lemma 1.5 Suppose that $\mathcal{F}^o \neq \emptyset$. Then for all $K \in \mathbb{R}$ the set

$$\{(x, s) \in \mathbb{R}^n \times \mathbb{R}^n \mid \exists y \in \mathbb{R}^m \text{ such that } (x, y, s) \in \mathcal{F}, s^T x \leq K\}$$

is bounded.

Proof. Let $(\bar{x}, \bar{y}, \bar{s}) \in \mathcal{F}^o$. For any $(x, y, s) \in \mathcal{F}$ we have $A\bar{x} = b$ and $Ax = b$, hence $A(\bar{x} - x) = 0$. Similarly, $A^T(\bar{y} - y) + (\bar{s} - s) = 0$. This implies

$$(\bar{s} - s)^T(\bar{x} - x) = -(\bar{y} - y)^T A(\bar{x} - x) = 0.$$

Assuming $s^T x \leq K$ we get

$$s^T \bar{x} + s^T x \leq K + \bar{s}^T \bar{x}.$$ 

The quantity $\xi := \min_i \min \{\bar{x}_i, \bar{s}_i\}$ is positive by assumption. We therefore get

$$\xi e^T (x + s) \leq K + \bar{s}^T \bar{x},$$

hence $\xi^{-1}(K + s^T \bar{x})$ is an upper bound on $x_i$ and $s_i$ for all $i$. \hfill \Box

Fix $\mu > 0$ and consider the barrier function

$$f: \mathcal{H}^o \to \mathbb{R}, \quad f(x, s) = \frac{1}{\mu} s^T x - \sum_{j=1}^n \ln(x_j s_j)$$ (1.9)
defined on the projection $\mathcal{H}^o$ of $\mathcal{F}^o$:

$$\mathcal{H}^o := \{ (x, s) \in \mathbb{R}^n \times \mathbb{R}^n \mid \exists y \in \mathbb{R}^m \ (x, y, s) \in \mathcal{F}^o \}.$$  

Note that $\mathcal{H}^o$ is convex as $\mathcal{F}^o$ is convex. Moreover, $f(x, s)$ approaches $\infty$ whenever any of the products $x_js_j$ approaches zero.

**Lemma 1.6**

1. $f$ is strictly convex.
2. $f$ is bounded from below.
3. For all $\kappa \in \mathbb{R}$ there exist $0 < \alpha < \beta$ such that

$$\{ (x, s) \in \mathcal{H}^o \mid f(x, s) \leq \kappa \} \subseteq [\alpha, \beta]^n \times [\alpha, \beta]^n.$$  

**Proof.** (1) Consider the function $g: \mathbb{R}_+^n \times \mathbb{R}_+^n \to \mathbb{R}$, $g(x, s) = -\sum_{j=1}^{n} \ln(x_js_j)$. We have $\frac{\partial^2 g}{\partial x^2}_j = x_{-2}^j$, $\frac{\partial^2 g}{\partial s^2}_j = s_{-2}^j$ and all other second order derivatives of $g$ vanish. The Hessian of $g$ is therefore positive definite and hence $g$ is strictly convex (cf. [2]). In particular, the restriction of $g$ to $\mathcal{H}^o$ is strictly convex as well.

We claim that the restriction of $s^T x$ to $\mathcal{H}^o$ is linear. In fact, if $\bar{x}$ is such that $A\bar{x} = b$, then for any $(x, s) \in \mathcal{H}^o$ with $y \in \mathbb{R}^m$ such that $(x, y, s) \in \mathcal{F}^o$

$$s^T x = c^T x - b^T y = c^T x - \bar{x}^T (c - s) = c^T x + \bar{x}^T s - \bar{x}^T c,$$

which is linear in $(x, s)$. This proves the first assumption.

(2) We write

$$f(x, s) = \sum_{j=1}^{n} h\left(\frac{x_js_j}{\mu}\right) + n - n \ln \mu,$$

where

$$h(t) := t - \ln t - 1.$$  

It is clear that $h$ is strictly convex on $(0, \infty)$ and $\lim_{t \to 0} h(t) = \infty$, $\lim_{t \to \infty} h(t) = \infty$. Moreover, $h(t) \geq 0$ for $t \in (0, \infty)$ with equality iff $t = 1$. Using this, we get

$$f(x, s) \geq n - n \ln \mu,$$

which shows the second assertion.

(3) Suppose $(x, s) \in \mathcal{H}^o$ with $f(x, s) \leq \kappa$ for some $\kappa$. Then, for all $j$,

$$h\left(\frac{x_js_j}{\mu}\right) \leq \kappa - n + n \log \mu =: \tilde{\kappa}.$$  

From the properties of $h$ it follows that there exist $0 < \alpha_1 < \beta_1$ such that

$$h^{-1}(\kappa, \tilde{\kappa}) \subseteq [\alpha_1, \beta_1],$$

hence $\mu \alpha_1 \leq x_js_j \leq \mu \beta_1$. Applying Lemma 1.5 with $K = n\mu \beta_1$ shows that there is some $\beta$ such that $x_j \leq \beta, s_j \leq \beta$. Hence $x_j \geq \mu \alpha_1 \beta^{-1}, s_j \geq \mu \alpha_1 \beta^{-1}$, which proves the third assertion with $\alpha = \mu \alpha_1 \beta^{-1}$. \qed
Complexity of the Linear Programming Feasibility Problem

Suppose that \( F^\circ \neq \emptyset \). Lemma 1.6(3) implies that \( f \) achieves its minimum in \( \mathcal{H}^\circ \). Moreover, the minimizer is unique as \( f \) is strictly convex. We shall denote this minimizer by \((x^\mu, s^\mu)\). We note that if \( \text{rk} A = m \leq n \), then \( y^\mu \) is uniquely determined by the condition \( A^Ty^\mu + s^\mu = c \).

To complete the argument, we will show that \( x_i, s_i = \mu, i = 1, 2, \ldots, n \), are exactly the first order conditions characterizing local minima of the function \( f \). (Note that a local minimum of \( f \) is a global minimum by the strict convexity of \( f \).)

We recall a well known fact about Lagrangian multipliers from analysis.

Let \( g, h_1, \ldots, h_m : U \to \mathbb{R} \) be differentiable functions defined on the open subset \( U \subseteq \mathbb{R}^n \). Suppose that \( u \in U \) is a local minimum of \( g \) under the constraints \( h_1 = 0, \ldots, h_m = 0 \). Then, if the gradients \( \nabla h_1, \ldots, \nabla h_m \) are linearly independent at \( u \), there exist Lagrange multipliers \( \lambda_1, \ldots, \lambda_m \in \mathbb{R} \) such that

\[
\nabla g(u) + \lambda_1 \nabla h_1(u) + \ldots + \lambda_m \nabla h_m(u) = 0.
\]

We apply this fact to the problem

\[
\begin{align*}
\min f(x, s) \quad & \text{s.t. } Ax = b, \ A^Ty + s = c, \ x > 0, \ s > 0.
\end{align*}
\]

Suppose that \((x, y, s)\) is a local minimum of \( f \). The linear independence condition holds due to Lemma 1.4 and our assumption \( \text{rk} A = m \leq n \). By (1.10) there are Lagrange multipliers \( v \in \mathbb{R}^m, w \in \mathbb{R}^n \) such that

\[
\mu^{-1}s - X^{-1}e + A^Tv = 0, \ Aw = 0, \ \mu^{-1}x - S^{-1}e + w = 0.
\]

(Here we have used that \( \frac{\partial f}{\partial x} = \mu^{-1}s - X^{-1}e, \ \frac{\partial f}{\partial y} = 0, \ \frac{\partial f}{\partial s} = \mu^{-1}x - S^{-1}e \).) The last two conditions imply

\[
A(\mu^{-1}x - S^{-1}e) = 0.
\]

Hence \((\mu^{-1}Xe - S^{-1}e)^TA^Tv = 0\). With the first condition we get

\[
(\mu^{-1}Xe - S^{-1}e)^T(\mu^{-1}Se - X^{-1}e) = 0.
\]

Therefore

\[
0 = (\mu^{-1}Xe - S^{-1}e)^T(X^{-1/2}S^{1/2})(X^{1/2}S^{-1/2})(\mu^{-1}Se - X^{-1}e) = \|\mu^{-1}(XS)^{1/2}e - (XS)^{-1/2}e\|^2.
\]

This implies \( XSe = \mu e \), hence \((x, y, s)\) lies on the central path \( C \).

Conversely, suppose that \((x, y, s) \in F^\circ\) satisfies \( XSe = \mu e \). Put \( v = 0 \) and \( w = 0 \). Then the first order conditions (1.11) are satisfied. Since \( f \) is strictly convex, \((x, s)\) is a global minimum of \( f \). By the previously shown uniqueness, we have \((x, s) = (x^\mu, s^\mu)\). This completes the proof of Theorem 1.2. \( \square \)
1.4 Analysis of IPM

Recall the following useful conventions: For a vector $u \in \mathbb{R}^d$ we denote by $U$ the matrix $U = \text{diag}(u_1, \ldots, u_d)$. Moreover, $e$ stands for the vector $(1, \ldots, 1)^T$ of the corresponding dimension. Note that $Ue = u$. The euclidean norm of $u$ is denoted by $\|u\|$. We write $\|U\|$ for the operator norm and $\|U\|_F$ for the Frobenius norm of $U$.

**Lemma 1.7** Let $u, v \in \mathbb{R}^d$ be such that $u^Tv \geq 0$. Then

$$\|UVe\| \leq \frac{1}{2}\|u + v\|^2.$$  

**Proof.** We have

$$\|UVe\| = \|Uv\| \leq \|U\| \|v\| \leq \|U\|_F \|v\| = \|u\| \|v\|.$$  

Moreover, as $u^Tv \geq 0$,

$$\|u\| \|v\| \leq \frac{1}{2}(\|u\|^2 + \|v\|^2) \leq \frac{1}{2}(\|u\|^2 + 2u^Tv + \|v\|^2) = \frac{1}{2}\|u + v\|^2.$$  

We remark that with a little more work, the factor $2^{-1}$ in Lemma 1.7 can be improved to $2^{-3/2}$, see [3, Lemma 14.1].

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ such that $\text{rk} A = m \leq n$. Moreover, let $z = (x, y, s) \in \mathcal{F}$, that is

$$Ax = b, \quad A^Ty + s = c, \quad x > 0, \quad s > 0.$$  

We consider one step of Algorithm 1.1, the primal-dual IPM, with centering parameter $\sigma \in (0, 1)$. That is, we set $\mu := \mu(z) = \frac{1}{n}s^Tx$, we define $\Delta z = (\Delta x, \Delta y, \Delta s)$ by

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sigma \mu e - XSe \end{bmatrix},$$

and put

$$\tilde{z} = (\tilde{x}, \tilde{y}, \tilde{s}) = (x, y, s) + (\Delta x, \Delta y, \Delta s).$$

**Lemma 1.8** 1. $\Delta s^T \Delta x = 0$

2. $\mu(\tilde{z}) = \sigma \mu(z)$

3. $\tilde{z} \in \mathcal{F}$ if $\tilde{x} > 0, \tilde{s} > 0$. 


Proof. (1) By definition of $\Delta z = (\Delta x, \Delta y, \Delta s)$ we have
\begin{align}
A^T \Delta y + \Delta s &= 0 \\
A \Delta x &= 0 \\
S \Delta x + X \Delta s &= \sigma \mu e - XSe
\end{align}
(1.13)

Therefore,
$$\Delta s^T \Delta x = -\Delta y^T A \Delta x = 0.$$  

(2) The third equation in (1.13) implies $s^T \Delta x + x^T \Delta s = n \sigma \mu - x^T s$. Therefore,
$$\tilde{s}^T \tilde{x} = (s^T + \Delta s^T)(x + \Delta x) = s^T x + \Delta s^T x + s^T \Delta x + \Delta s^T \Delta x = n \sigma \mu.$$  

This means $\mu(\tilde{z}) = \frac{1}{n} \tilde{s}^T \tilde{x} = \sigma \mu$.

(3) We already verified at the end of Section 1.2 (by a straightforward calculation) that $\tilde{z}$ satisfies the equality constraints in (1.5). 

A remaining issue is how to achieve that $\tilde{x} > 0$, $\tilde{s} > 0$ by a suitable choice of the centering parameter $\sigma$.

**Definition 1.9** Let $\beta > 0$. The central neighborhood $\mathcal{N}(\beta)$ is defined as the set of strictly feasible points $z = (x, y, s) \in \mathcal{F}^\circ$ such that
$$\|XSe - \mu(z)e\| \leq \beta \mu(z).$$

The central neighborhood is a neighborhood of the central path $C$ in $\mathcal{F}^\circ$ that becomes smaller when $\mu(z)$ approaches zero (cf. Figure 1.1).

In the following we set $\beta = \frac{1}{4}$ and write $\mathcal{N} := \mathcal{N}(\frac{1}{4})$.

**Lemma 1.10** Let $z = (x, y, s) \in \mathcal{N}$ and $\Delta z = (\Delta x, \Delta y, \Delta s)$ be defined by (1.12) with respect to $\sigma = 1 - \frac{\xi}{\sqrt{n}}$ with $0 \leq \xi \leq \frac{1}{4}$. Then $\tilde{z} = z + \Delta z$ satisfies $\tilde{z} \in \mathcal{N}$.

**Proof.** By (1.12) we have
\begin{align}
XSe + X \Delta s + S \Delta x &= \sigma \mu e, \\
\end{align}
(1.14) which implies
$$\tilde{X} \tilde{S}e = XSe + X \Delta s + S \Delta x + \Delta X \Delta S e = \Delta X \Delta S e + \sigma \mu e.$$  

Moreover, by Lemma 1.8(2), $\mu(\tilde{z}) = \sigma \mu$. We therefore need to show that
\begin{align}
\|\tilde{X} \tilde{S}e - \sigma \mu e\| = \|\Delta X \Delta S e\| \leq \beta \mu(\tilde{z}) = \beta \sigma \mu.
\end{align}
(1.15)
To do so note first that $z \in \mathcal{N}$ implies $|x_i s_i - \mu| \leq \beta \mu$ for all $i$, hence
\begin{equation}
(1 - \beta) \mu \leq x_i s_i \leq (1 + \beta) \mu.
\end{equation}
By (1.14) we have $X \Delta s + s \Delta x = \sigma \mu e - X Se$. Setting $D := X^{1/2} S^{-1/2}$ we get
\begin{equation}
D \Delta s + D^{-1} \Delta x = (XS)^{-1/2}(\sigma \mu e - X Se).
\end{equation}
Because $(D^{-1} \Delta x)^T (D \Delta s) = \Delta s^T \Delta x = 0$ (cf. Lemma 1.8(1)) we can apply Lemma 1.7 with $u = D^{-1} \Delta x$ and $v = D \Delta s$ to obtain
\begin{align*}
\|X \Delta Se\| & = \|(D^{-1} X)(D \Delta S)e\| \\
& \leq 2^{-1} \|D^{-1} \Delta x + D \Delta s\|^2 \\
& \leq 2^{-1} \|X\|^{-1/2} (\sigma \mu e - X Se)\|^2 \\
& \leq (2(1 - \beta))^{-1} \|\sigma \mu e - X Se\|^2 \\
& \leq (2(1 - \beta))^{-1} (\|\mu e - X Se\| + \|\mu(\sigma - 1)e\|)^2 \\
& \leq (2(1 - \beta))^{-1} (\beta \mu + (1 - \sigma) \mu) e)^2 \\
& \leq (2(1 - \beta))^{-1} (\beta + \xi)^2 \mu \\
& \text{by def. of } \sigma.
\end{align*}
A small calculation shows that
\begin{equation}
\frac{1}{2(1 - \beta)} (\beta + \xi)^2 \leq \beta (1 - \xi) \leq \beta (1 - \frac{\xi}{\sqrt{n}})
\end{equation}
for $\beta = \frac{1}{4}$ and $0 \leq \xi \leq \frac{1}{4}$. This proves (1.15).

We still need to show that $\tilde{z} \in \mathcal{F}^\circ$. For this, by Lemma 1.8(3), it is sufficient to prove that $\tilde{x} \tilde{s} > 0$. Inequality (1.15) implies $\tilde{x}_i \tilde{s}_i \geq (1 - \beta) \sigma \mu > 0$. Suppose we had $\tilde{x}_i < 0$ or $\tilde{s}_i < 0$ for some $i$. Then $\tilde{x}_i < 0$ and $\tilde{s}_i < 0$ which implies $|\Delta x_i| > x_i$ and $|\Delta s_i| > s_i$. But then, $\beta \mu > \beta \sigma \mu \geq \|X \Delta Se\| \geq |\Delta x_i \Delta s_i| > x_i s_i \geq (1 - \beta) \mu$, hence $\beta \geq \frac{1}{4}$, a contradiction.

\begin{proof}

Theorem 1.11 On an input $(\mathcal{A}, b, c) \in \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n$ with \( \text{rk} \, \mathcal{A} = m \leq n \) and the choice of the centering parameter $\sigma = 1 - \frac{\xi}{\sqrt{n}}$ with $\xi \in (0, \frac{1}{4}]$, Algorithm 1.1 produces on a strictly feasible starting point $z_0$ in the central neighborhood $\mathcal{N} = \mathcal{N}(\frac{1}{4})$ a sequence of iterates $z_k \in \mathcal{N}$ such that $\mu(z_k) = \sigma^k \mu(z_0)$, for $k \in \mathbb{N}$. We therefore have for all $\varepsilon > 0$,
\[
\mu(z_k) \leq \varepsilon \quad \text{for} \quad k \geq \frac{\sqrt{n}}{\xi} \ln \frac{\mu(z_0)}{\varepsilon}.
\]
\end{proof}

Proof. It suffices to show the last assertion. This follows from the implication
\[
k \geq a^{-1} \ln B \Rightarrow (1 - a)^k \leq \frac{1}{B}
\]
for $0 < a < 1$, $B > 0$. (Use $\ln(1 - a) \leq -a$ to show this.)

\end{proof}
Bibliography


