Primal-Dual Interior-Point Methods for Linear Programming based on Newton's Method

Robert M. Freund

March, 2004

1 The Problem

The logarithmic barrier approach to solving a linear program dates back to the work of Fiacco and McCormick in 1967 in their book *Sequential Unconstrained Minimization Techniques*, also known simply as SUMT. The method was not believed then to be either practically or theoretically interesting, when in fact today it is both! The method was re-born as a consequence of Karmarkar's interior-point method, and has been the subject of an enormous amount of research and computation, even to this day. In these notes we present the basic algorithm and a basic analysis of its performance.

Consider the linear programming problem in standard form:

$$P: \quad \text{minimize} \quad c^T x$$

$$\text{s.t.} \quad Ax = b$$
$$x \geq 0,$$

where x is a vector of n variables, whose standard linear programming dual problem is:

$$D: \quad \text{maximize} \quad b^T \pi$$

$$\text{s.t.} \quad A^T \pi + s = c$$
$$s \geq 0.$$

Given a feasible solution x of P and a feasible solution (π, s) of D, the duality gap is simply

$$c^T x - b^T \pi = x^T s \geq 0.$$

We introduce the following notation which will be very convenient for manipulating equations, etc. Suppose that $x > 0$. Define the matrix X to be the $n \times n$ diagonal matrix whose diagonal entries are precisely the components of x. Then X looks like:

$$\begin{pmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x_n \end{pmatrix}.$$

Notice that X is positive definite, and so is X^2, which looks like:

$$\begin{pmatrix} (x_1)^2 & 0 & \ldots & 0 \\ 0 & (x_2)^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & (x_n)^2 \end{pmatrix}.$$

Similarly, the matrices X^{-1} and X^{-2} look like:

$$\begin{pmatrix} (1/x_1) & 0 & \ldots & 0 \\ 0 & (1/x_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & (1/x_n) \end{pmatrix}$$

and

$$\begin{pmatrix} 1/(x_1)^2 & 0 & \dots & 0 \\ 0 & 1/(x_2)^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/(x_n)^2 \end{pmatrix}.$$

Let us introduce a logarithmic barrier term for P. We obtain:

$$P(\theta): \quad \text{minimize} \quad c^T x - \theta \sum_{j=1}^{n} \ln(x_j)$$

$$\text{s.t.} \qquad Ax = b$$
$$x > 0.$$

Because the gradient of the objective function of $P(\theta)$ is simply $c - \theta X^{-1} e$, (where e is the vector of ones, i.e., $e = (1, 1, 1, ..., 1)^T$), the Karush-Kuhn-Tucker conditions for $P(\theta)$ are:

$$\begin{cases} Ax = b, x > 0 \\ \\ c - \theta X^{-1} e = A^T \pi \ . \end{cases} \tag{1}$$

If we define $s = \theta X^{-1} e$, then

$$\frac{1}{\theta} X s = e \ ,$$

4

equivalently

$$\frac{1}{\theta} X S e = e \ ,$$

and we can rewrite the Karush-Kuhn-Tucker conditions as:

$$\begin{cases} Ax = b, x > 0 \\ A^T \pi + s = c \\ \frac{1}{\theta} X S e - e = 0 \ . \end{cases} \tag{2}$$

From the equations of (2) it follows that if (x, π, s) is a solution of (2), then x is feasible for P, (π, s) is feasible for D, and the resulting duality gap is:

$$x^T s = e^T X S e = \theta e^T e = \theta n.$$

This suggests that we try solving $P(\theta)$ for a variety of values of θ as $\theta \to 0$.

However, we cannot usually solve (2) exactly, because the third equation group is not linear in the variables. We will instead define a "β-approximate solution" of the Karush-Kuhn-Tucker conditions (2). A "β-approximate solution" of $P(\theta)$ is defined as any solution (x, π, s) of

$$\begin{cases} Ax = b, x > 0 \\ A^T \pi + s = c \\ \|\frac{1}{\theta} X s - e\| \le \beta \ . \end{cases} \tag{3}$$

Here the norm $\| \cdot \|$ is the Euclidean norm.

Lemma 1.1 *If $(\bar{x}, \bar{\pi}, \bar{s})$ is a β-approximate solution of $P(\theta)$ and $\beta < 1$, then $\bar{x}$ is feasible for P, $(\bar{\pi}, \bar{s})$ is feasible for D, and the duality gap satisfies:*

$$n\theta(1-\beta) \leq c^T\bar{x} - b^T\bar{\pi} = \bar{x}^T\bar{s} \leq n\theta(1+\beta). \tag{4}$$

Proof: Primal feasibility is obvious. To prove dual feasibility, we need to show that $\bar{s} \geq 0$. To see this, note that the third equation system of (3) implies that

$$-\beta \leq \frac{\bar{x}_j\bar{s}_j}{\theta} - 1 \leq \beta$$

which we can rearrange as:

$$\theta(1-\beta) \leq \bar{x}_j\bar{s}_j \leq \theta(1+\beta). \tag{5}$$

Therefore $\bar{x}_j\bar{s}_j \geq (1-\beta)\theta > 0$, which implies $\bar{x}_j\bar{s}_j > 0$, and so $\bar{s}_j > 0$. From (5) we have

$$n\theta(1-\beta) = \sum_{j=1}^{n}\theta(1-\beta) \leq \sum_{j=1}^{n}\bar{x}_j\bar{s}_j = \bar{x}^T\bar{s} \leq \sum_{j=1}^{n}\theta(1+\beta) = n\theta(1+\beta).$$

∎

2   The Primal-Dual Algorithm

Based on the analysis just presented, we are motivated to develop the following algorithm:

Step 0 . Initialization. Data is $(x^0, \pi^0, s^0, \theta^0)$. $k = 0$. Assume that (x^0, π^0, s^0) is a β-approximate solution of $P(\theta^0)$ for some known value of β that satisfies $\beta < \frac{1}{2}$.

Step 1. Set current values. $(\bar{x}, \bar{\pi}, \bar{s}) = (x^k, \pi^k, s^k)$, $\theta = \theta^k$.

Step 2. Shrink θ. Set $\theta' = \alpha\theta$ for some $\alpha \in (0, 1)$.

Step 3. Compute the primal-dual Newton direction. Compute the Newton step $(\Delta x, \Delta \pi, \Delta s)$ for the equation system (2) at $(x, \pi, s) = (\bar{x}, \bar{\pi}, \bar{s})$ for $\theta = \theta'$, by solving the following system of equations in the variables $(\Delta x, \Delta \pi, \Delta s)$:

$$\begin{cases} A\Delta x = 0 \\\\ A^T\Delta\pi + \Delta s = 0 \\\\ \bar{S}\Delta x + \bar{X}\Delta s = \bar{X}\bar{S}e - \theta e \; . \end{cases} \qquad (6)$$

Denote the solution to this system by $(\Delta x, \Delta \pi, \Delta s)$.

Step 4. Update All Values.

$$(x', \pi', s') = (\bar{x}, \bar{\pi}, \bar{s}) + (\Delta x, \Delta \pi, \Delta s)$$

Step 5. Reset Counter and Continue. $(x^{k+1}, \pi^{k+1}, s^{k+1}) = (x', \pi', s')$. $\theta^{k+1} = \theta'$. $k \leftarrow k + 1$. Go to Step 1.

Figure 1 shows a picture of the algorithm.

Some of the issues regarding this algorithm include:

- how to set the approximation constant β and the fractional decrease parameter α. We will see that it will be convenient to set $\beta = \frac{3}{40}$ and

$$\alpha = 1 - \frac{\frac{1}{8}}{\frac{1}{5} + \sqrt{n}} \; .$$

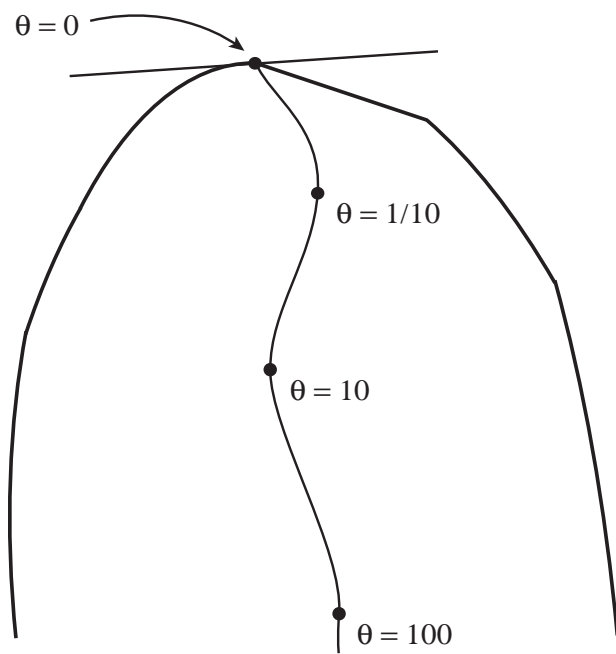- the derivation of the primal-dual Newton equation system (6)

7

Figure 1: A conceptual picture of the interior-point algorithm.

- whether or not successive iterative values (x^k, π^k, s^k) are β-approximate solutions to $P(\theta^k)$

- how to get the method started

3 The Primal-Dual Newton Step

We introduce a logarithmic barrier term for P to obtain $P(\theta)$:

$$P(\theta): \quad \text{minimize}_x \quad c^T x - \theta \sum_{j=1}^{n} \ln(x_j)$$
$$\text{s.t.} \quad Ax = b$$
$$x > 0.$$

Because the gradient of the objective function of $P(\theta)$ is simply $c - \theta X^{-1}e$, (where e is the vector of ones, i.e., $e = (1, 1, 1, ..., 1)^T$), the Karush-Kuhn-Tucker conditions for $P(\theta)$ are:

$$\begin{cases} Ax = b, x > 0 \\ \\ c - \theta X^{-1}e = A^T \pi \ . \end{cases} \tag{7}$$

We define $s = \theta X^{-1}e$, whereby

$$\frac{1}{\theta} Xs = e \ ,$$

equivalently

$$\frac{1}{\theta} XSe = e \ ,$$

and we can rewrite the Karush-Kuhn-Tucker conditions as:

$$\begin{cases} Ax = b, x > 0 \\ A^T\pi + s = c \\ \frac{1}{\theta}XSe - e = 0 \ . \end{cases} \tag{8}$$

Let $(\bar{x}, \bar{\pi}, \bar{s})$ be our current iterate, which we assume is primal and dual feasible, namely:

$$A\bar{x} = b, \ \bar{x} > 0 \ , \quad A^T\bar{\pi} + \bar{s} = c \ , \bar{s} > 0 \ . \tag{9}$$

Introducing a direction $(\Delta x, \Delta\pi, \Delta s)$, the next iterate will be $(\bar{x}, \bar{\pi}, \bar{s}) + (\Delta x, \Delta\pi, \Delta s)$, and we want to solve:

$$\begin{cases} A(\bar{x} + \Delta x) = b, \bar{x} + \Delta x > 0 \\ A^T(\bar{\pi} + \Delta\pi) + (\bar{s} + \Delta s) = c \\ \frac{1}{\theta}(\bar{X} + \Delta X)(\bar{S} + \Delta S)e - e = 0 \ . \end{cases}$$

Keeping in mind that $(\bar{x}, \bar{\pi}, \bar{s})$ is primal-dual feasible and so satisfies (9), we can rearrange the above to be:

$$\begin{cases} A\Delta x = 0 \\ A^T\Delta\pi + \Delta s = 0 \\ \bar{S}\Delta x + \bar{X}\Delta s = \theta e - \bar{X}\bar{S}e - \Delta X\Delta Se \ . \end{cases}$$

Notice that the only nonlinear term in the above system of equations in $(\Delta x, \Delta\pi, \Delta s)$ is term the term "$\Delta X\Delta Se$" in the last system. If we erase this term, which is the same as the linearized version of the equations, we obtain the following primal-dual Newton equation system:

$$\begin{cases} A\Delta x = 0 \\[2mm] A^T \Delta \pi + \Delta s = 0 \\[2mm] \bar{S}\Delta x + \bar{X}\Delta s = \theta e - \bar{X}\bar{S}e \ . \end{cases} \quad (10)$$

The solution $(\Delta x, \Delta \pi, \Delta s)$ of the system (10) is called the primal-dual Newton step. We can manipulate these equations to yield the following formulas for the solution:

$$\begin{aligned} \Delta x \ &\leftarrow \ \left[I - \bar{X}\bar{S}^{-1}A^T \left(A\bar{X}\bar{S}^{-1}A^T \right)^{-1} A \right] \left(-\bar{x} + \theta \bar{S}^{-1}e \right) \ , \\[3mm] \Delta \pi \ &\leftarrow \ \left(A\bar{X}\bar{S}^{-1}A^T \right)^{-1} A \left(\bar{x} - \theta \bar{S}^{-1}e \right) \ , \\[3mm] \Delta s \ &\leftarrow \ A^T \left(A\bar{X}\bar{S}^{-1}A^T \right)^{-1} A \left(-\bar{x} + \theta \bar{S}^{-1}e \right) \ . \end{aligned} \quad (11)$$

Notice, by the way, that the computational effort in these equations lies primarily in solving a single equation:

$$\left(A\bar{X}\bar{S}^{-1}A^T \right) \Delta \pi = A \left(\bar{x} - \theta \bar{S}^{-1}e \right) \ .$$

Once this system is solved, we can easily substitute:

$$\begin{aligned} \Delta s \ &\leftarrow \ -A^T \Delta \pi \\[3mm] \Delta x \ &\leftarrow \ -\bar{x} + \theta \bar{S}^{-1}e - \bar{S}^{-1}\bar{X}\Delta s \ . \end{aligned} \quad (12)$$

However, let us instead simply work with the primal-dual Newton system (10). Suppose that $(\Delta x, \Delta \pi, \Delta s)$ is the (unique) solution of the primal-dual Newton system (10). We obtain the new value of the variables (x, π, s) by taking the Newton step:

$$(x^{'}, \pi^{'}, s^{'}) = (\bar{x}, \bar{\pi}, \bar{s}) + (\Delta x, \Delta \pi, \Delta s) \ .$$

We have the following very powerful convergence theorem which demonstrates the quadratic convergence of Newton's method for this problem, with an explicit guarantee of the range in which quadratic convergence takes place.

Theorem 3.1 (Explicit Quadratic Convergence of Newton's Method).
Suppose that $(\bar{x}, \bar{\pi}, \bar{s})$ is a β-approximate solution of $P(\theta)$ and $\beta < \frac{1}{2}$. Let $(\Delta x, \Delta \pi, \Delta s)$ be the solution to the primal-dual Newton equations (10), and let:
$$(x^{'}, \pi^{'}, s^{'}) = (\bar{x}, \bar{\pi}, \bar{s}) + (\Delta x, \Delta \pi, \Delta s) \ .$$
Then $(x^{'}, \pi^{'}, s^{'})$ is a $\left(\frac{1+\beta}{(1-\beta)^2} \right) \beta^2$-approximate solution of $P(\theta)$.

Proof: Our current point $(\bar{x}, \bar{\pi}, \bar{s})$ satisfies:

$$\begin{cases} A\bar{x} = b, \bar{x} > 0 \\[2mm] A^T \bar{\pi} + \bar{s} = c \\[2mm] \| \frac{1}{\theta} \bar{X} \bar{S} e - e \| \leq \beta \ . \end{cases} \quad (13)$$

Furthermore the primal-dual Newton step $(\Delta x, \Delta \pi, \Delta s)$ satisfies:

$$\begin{cases} A\Delta x = 0 \\[2mm] A^T \Delta \pi + \Delta s = 0 \\[2mm] \bar{S}\Delta x + \bar{X}\Delta s = \theta e - \bar{X}\bar{S}e \ . \end{cases} \quad (14)$$

Note from the first two equations of (14) that $\Delta x^T \Delta s = 0$. From the third equation of (13) we have

$$1 - \beta \leq \frac{\bar{s}_j \bar{x}_j}{\theta} \leq 1 + \beta, \ j = 1, \ldots, n, \quad (15)$$

which implies:

$$\bar{x}_j \geq \frac{(1-\beta)\theta}{\bar{s}_j} \text{ and } \bar{s}_j \geq \frac{(1-\beta)\theta}{\bar{x}_j} \, , \, j = 1, \ldots, n \, . \tag{16}$$

As a result of this we obtain:

$$
\begin{aligned}
\theta(1-\beta)\|\bar{X}^{-1}\Delta x\|^2 \quad &= \theta(1-\beta)\Delta x^T \bar{X}^{-1}\bar{X}^{-1}\Delta x \\
&\leq \Delta x^T \bar{X}^{-1}\bar{S}\Delta x \\
&= \Delta x^T \bar{X}^{-1}\left(\theta e - \bar{X}\bar{S}e - \bar{X}\Delta s\right) \\
&= \Delta x^T \bar{X}^{-1}\left(\theta e - \bar{X}\bar{S}e\right) \\
&\leq \|\bar{X}^{-1}\Delta x\|\|\theta e - \bar{X}\bar{S}e\| \\
&\leq \|\bar{X}^{-1}\Delta x\|\beta\theta \, .
\end{aligned}
$$

From this it follows that

$$\|\bar{X}^{-1}\Delta x\| \leq \frac{1-\beta}{\beta} < 1 \, .$$

Therefore

$$x^{'} = \bar{x} + \Delta x = \bar{X}(e + \bar{X}^{-1}\Delta x) > 0 \, .$$

We have the exact same chain of inequalities for the dual variables:

$$
\begin{aligned}
\theta(1-\beta)\|\bar{S}^{-1}\Delta s\|^2 \quad &= \theta(1-\beta)\Delta s^T \bar{S}^{-1}\bar{S}^{-1}\Delta s \\
&\leq \Delta s^T \bar{S}^{-1}\bar{X}\Delta s \\
&= \Delta s^T \bar{S}^{-1}\left(\theta e - \bar{X}\bar{S}e - \bar{S}\Delta x\right) \\
&= \Delta s^T \bar{S}^{-1}\left(\theta e - \bar{X}\bar{S}e\right) \\
&\leq \|\bar{S}^{-1}\Delta s\|\|\theta e - \bar{X}\bar{S}e\| \\
&\leq \|\bar{S}^{-1}\Delta s\|\beta\theta \, .
\end{aligned}
$$

From this it follows that

$$\|\bar{S}^{-1}\Delta s\| \leq \frac{1-\beta}{\beta} < 1 \, .$$

Therefore

$$s^{'} = \bar{s} + \Delta s = \bar{S}(e + \bar{S}^{-1}\Delta s) > 0 \, .$$

Next note from (14) that for $j = 1, \ldots, n$ we have:

$$x^{'}_j s^{'}_j = (\bar{x}_j + \Delta x_j)(\bar{s}_j + \Delta s_j) = \bar{x}_j\bar{s}_j + \bar{x}_j\Delta s_j + \Delta x_j\bar{s}_j + \Delta x_j\Delta s_j = \theta + \Delta x_j\Delta s_j \, .$$

13

Therefore

$$\left(e - \frac{1}{\theta}X'S'e\right)_j = -\frac{\Delta x_j \Delta s_j}{\theta} \ .$$

From this we obtain:

$$\begin{aligned}
\|e - \tfrac{1}{\theta}X'S'e\| &\leq \|e - \tfrac{1}{\theta}X'S'e\|_1 \\
&= \sum_{j=1}^{n} \frac{|\Delta x_j \Delta s_j|}{\theta} \\
&= \sum_{j=1}^{n} \frac{|\Delta x_j|}{\bar{x}_j} \frac{\Delta s_j}{\bar{s}_j} \frac{\bar{x}_j \bar{s}_j}{\theta} \\
&\leq \sum_{j=1}^{n} \frac{|\Delta x_j|}{\bar{x}_j} \frac{|\Delta s_j|}{\bar{s}_j}(1+\beta) \\
&\leq \|\bar{X}^{-1}\Delta x\|\|\bar{S}^{-1}\Delta s\|(1+\beta) \\
&\leq \left(\frac{\beta}{1-\beta}\right)^2 (1+\beta)
\end{aligned}$$

∎

4   Complexity Analysis of the Algorithm

Theorem 4.1 (Relaxation Theorem). *Suppose that* $(\bar{x}, \bar{\pi}, \bar{s})$ *is a* $\beta = \frac{3}{40}$-*approximate solution of* $P(\theta)$. *Let*

$$\alpha = 1 - \frac{\frac{1}{8}}{\frac{1}{5} + \sqrt{n}}$$

and let $\theta' = \alpha\theta$. *Then* $(\bar{x}, \bar{\pi}, \bar{s})$ *is a* $\beta = \frac{1}{5}$-*approximate solution of* $P(\theta')$.

Proof: The triplet $(\bar{x}, \bar{\pi}, \bar{s})$ satisfies $A\bar{x} = b, \bar{x} > 0$, and $A^T\bar{\pi} + \bar{s} = c$, and so it remains to show that

$$\left\|\frac{1}{\theta'}\bar{X}\bar{s} - e\right\| \leq \frac{1}{5} \ .$$

We have

$$\left\|\frac{1}{\theta'}\bar{X}\bar{s} - e\right\| = \left\|\frac{1}{\alpha\theta}\bar{X}\bar{s} - e\right\| = \left\|\frac{1}{\alpha}\left(\frac{1}{\theta}\bar{X}\bar{s} - e\right) - \left(1 - \frac{1}{\alpha}\right)e\right\|$$

$$\leq \left(\frac{1}{\alpha}\right)\left\|\frac{1}{\theta}\bar{X}\bar{s} - e\right\| + \left|\frac{1-\alpha}{\alpha}\right|\|e\|$$

14

$$\leq \frac{\frac{3}{40}}{\alpha} + \left(\frac{1-\alpha}{\alpha}\right)\sqrt{n} \quad = \quad \frac{\frac{3}{40}+\sqrt{n}}{\alpha} - \sqrt{n} = \frac{1}{5} \ .$$

∎

Theorem 4.2 (Convergence Theorem). *Suppose that (x^0, π^0, s^0) is a $\beta = \frac{3}{40}$-approximate solution of $P(\theta^0)$. Then for all $k = 1, 2, 3, \ldots$, (x^k, π^k, s^k) is a $\beta = \frac{3}{40}$-approximate solution of $P(\theta^k)$.*

Proof: By induction, suppose that the theorem is true for iterates $0, 1, 2, ..., k$. Then (x^k, π^k, s^k) is a $\beta = \frac{3}{40}$-approximate solution of $P(\theta^k)$. From the Relaxation Theorem, (x^k, π^k, s^k) is a $\frac{1}{5}$-approximate solution of $P(\theta^{k+1})$ where $\theta^{k+1} = \alpha\theta^k$.

From the Quadratic Convergence Theorem, $(x^{k+1}, \pi^{k+1}, s^{k+1})$ is a β-approximate solution of $P(\theta^{k+1})$ for

$$\beta = \frac{1+\frac{1}{5}}{\left(1-\frac{1}{5}\right)^2}\left(\frac{1}{5}\right)^2 = \frac{3}{40} \ .$$

Therefore, by induction, the theorem is true for all values of k. ∎

Figure 2 shows a better picture of the algorithm.

Theorem 4.3 (Complexity Theorem). *Suppose that (x^0, π^0, s^0) is a $\beta = \frac{3}{40}$-approximate solution of $P(\theta^0)$. In order to obtain primal and dual feasible solutions (x^k, π^k, s^k) with a duality gap of at most ϵ, one needs to run the algorithm for at most*

$$k = \left\lceil 10\sqrt{n}\ln\left(\frac{43}{37}\frac{(x^0)^T s^0}{\epsilon}\right)\right\rceil$$

iterations.

Proof: Let k be as defined above. Note that

$$\alpha = 1 - \frac{\frac{1}{8}}{\frac{1}{5}+\sqrt{n}} = 1 - \frac{1}{\left(\frac{8}{5}+8\sqrt{n}\right)} \leq 1 - \frac{1}{10\sqrt{n}}.$$
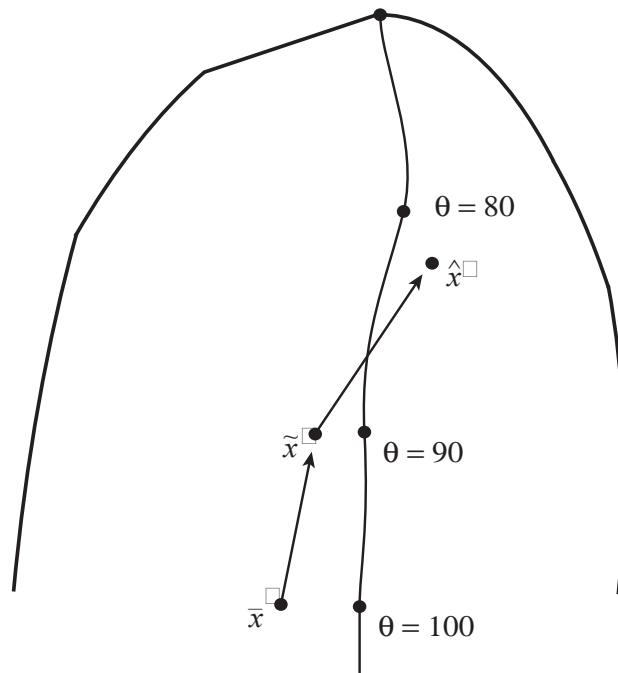
15

Figure 2: Another picture of the interior-point algorithm.

Therefore

$$\theta^k \le \left(1 - \frac{1}{10\sqrt{n}}\right)^k \theta^0.$$

This implies that

$$c^T x^k - b^T \pi^k = (x^k)^T s^k \le \theta^k n(1+\beta) \le \left(1 - \frac{1}{10\sqrt{n}}\right)^k \left(\frac{43}{40}n\right)\theta^0)$$

$$\le \left(1 - \frac{1}{10\sqrt{n}}\right)^k \left(\frac{43}{40}n\right)\left(\frac{(x^0)^T s^0}{\frac{37}{40}n}\right),$$

from (4). Taking logarithms, we obtain

$$\ln(c^T x^k - b^T \pi^k) \le k \ln\left(1 - \frac{1}{10\sqrt{n}}\right) + \ln\left(\frac{43}{37}(x^0)^T s^0\right)$$

$$\le \frac{-k}{10\sqrt{n}} + \ln\left(\frac{43}{37}(x^0)^T s^0\right)$$

$$\le -\ln\left(\frac{43}{37}\frac{(x^0)^T s^0}{\epsilon}\right) + \ln\left(\frac{43}{37}(x^0)^T s^0\right) = \ln(\epsilon).$$

Therefore $c^T x^k - b^T \pi^k \le \epsilon.$ ∎

5   An Implementable Primal-Dual Interior-Point Algorithm

Herein we describe a more implementable primal-dual interior-point algorithm. This algorithm differs from the previous method in the following respects:

- We do not assume that the current point is near the central path. In fact, we do not assume that the current point is even feasible.

- The fractional decrease parameter α is set to $\alpha = \frac{1}{10}$ rather than the conservative value of

$$\alpha = 1 - \frac{\frac{1}{8}}{\frac{1}{5} + \sqrt{n}}.$$

17

- We do not necessarily take the full Newton step at each iteration, and we take different step-sizes in the primal and dual.

Our current "point" is $(\bar{x}, \bar{\pi}, \bar{s})$ for which $\bar{x} > 0$ and $\bar{s} > 0$, but quite possibly $A\bar{x} \neq b$ and/or $A^T\bar{\pi} + \bar{s} \neq c$. We are given a value of the central path barrier parameter $\theta > 0$. We want to compute a direction $(\Delta x, \Delta \pi, \Delta s)$ so that $(\bar{x} + \Delta x, \bar{\pi} + \Delta \pi, \bar{s} + \Delta s)$ approximately solves the central path equations. We set up the system:

$$(1) \quad A(\bar{x} + \Delta x) = b$$

$$(2) \quad A^T(\bar{\pi} + \Delta \pi) + (\bar{s} + \Delta s) = c$$

$$(3) \quad (\bar{X} + \Delta X)(\bar{S} + \Delta S)e = \theta e \ .$$

We linearize this system of equations and rearrange the terms to obtain the Newton equations for the current point $(\bar{x}, \bar{\pi}, \bar{s})$:

$$(1) \quad A\Delta x = b - A\bar{x} =: r_1$$

$$(2) \quad A^T\Delta \pi + \Delta s = c - A^T\bar{\pi} - \bar{s} =: r_2$$

$$(3) \quad \bar{S}\Delta x + \bar{X}\Delta s = \theta e - \bar{X}\bar{S}e =: r_3$$

We refer to the solution $(\Delta x, \Delta \pi, \Delta s)$ to the above system as the primal-dual Newton direction at the point $(\bar{x}, \bar{\pi}, \bar{s})$. It differs from that derived earlier only in that earlier it was assumed that $r_1 = 0$ and $r_2 = 0$.

Given our current point $(\bar{x}, \bar{\pi}, \bar{s})$ and a given value of θ, we compute the Newton direction $(\Delta x, \Delta \pi, \Delta s)$ and we update our variables by choosing primal and dual step-sizes α_P and α_D to obtain new values:

$$(\tilde{x}, \tilde{\pi}, \tilde{s}) \leftarrow (\bar{x} + \alpha_P\Delta x, \bar{\pi} + \alpha_D\Delta \pi, \bar{s} + \alpha_D\Delta s) \ .$$

In order to ensure that $\tilde{x} > 0$ and $\tilde{s} > 0$, we choose a value of r satisfying $0 < r < 1$ ($r = 0.99$ is a common value in practice), and determine α_P and α_D as follows:

$$\alpha_P = \min\left\{1, \; r \min_{\Delta x_j < 0}\left\{\frac{\bar{x}_j}{-\Delta x_j}\right\}\right\}$$

$$\alpha_D = \min\left\{1, \; r \min_{\Delta s_j < 0}\left\{\frac{\bar{s}_j}{-\Delta s_j}\right\}\right\}.$$

These step-sizes ensure that the next iterate $(\tilde{x}, \tilde{\pi}, \tilde{s})$ satisfies $\tilde{x} > 0$ and $\tilde{s} > 0$.

5.1   Decreasing the Path Parameter θ

We also want to shrink θ at each iteration, in order to (hopefully) shrink the duality gap. The current iterate is $(\bar{x}, \bar{\pi}, \bar{s})$, and the current values satisfy:

$$\theta \approx \frac{\bar{x}^T \bar{s}}{n}$$

We then re-set θ to

$$\theta \leftarrow \left(\frac{1}{10}\right)\left(\frac{\bar{x}^T \bar{s}}{n}\right),$$

where the fractional decrease $\frac{1}{10}$ is user-specified.

5.2   The Stopping Criterion

We typically declare the problem solved when it is "almost" primal feasible, "almost" dual feasible, and there is "almost" no duality gap. We set our

tolerance ϵ to be a small positive number, typically $\epsilon = 10^{-8}$, for example, and we stop when:

$$(1) \quad \|A\bar{x} - b\| \le \epsilon$$

$$(2) \quad \|A^T\bar{\pi} + \bar{s} - c\| \le \epsilon$$

$$(3) \quad \bar{s}^T\bar{x} \le \epsilon$$

5.3   The Full Interior-Point Algorithm

1. Given (x^0, π^0, s^0) satisfying $x^0 > 0$, $s^0 > 0$, and $\theta^0 > 0$, and r satisfying $0 < r < 1$, and $\epsilon > 0$. Set $k \leftarrow 0$.

2. Test stopping criterion. Check if:

$$(1) \quad \|Ax^k - b\| \le \epsilon$$

$$(2) \quad \|A^T\pi^k + s^k - c\| \le \epsilon$$

$$(3) \quad (s^k)^Tx^k \le \epsilon \; .$$

   If so, STOP. If not, proceed.

3. Set $\theta \leftarrow \left(\dfrac{1}{10}\right)\left(\dfrac{(x^k)^T(s^k)}{n}\right)$

4. Solve the Newton equation system:

$$(1) \quad A\Delta x = b - Ax^k =: r_1$$

$$(2) \quad A^T\Delta\pi + \Delta s = c - A^T\pi^k - s^k =: r_2$$

$$(3) \quad S^k\Delta x + X^k\Delta s = \theta e - X^kS^ke =: r_3$$

20

5. Determine the step-sizes:

$$\theta_P = \min\left\{1,\; r\min_{\Delta x_j < 0}\left\{\frac{x_j^k}{-\Delta x_j}\right\}\right\}$$

$$\theta_D = \min\left\{1,\; r\min_{\Delta s_j < 0}\left\{\frac{s_j^k}{-\Delta s_j}\right\}\right\}.$$

6. Update values:

$$(x^{k+1}, \pi^{k+1}, s^{k+1}) \;\leftarrow\; (x^k + \alpha_P \Delta x,\, \pi^k + \alpha_D \Delta \pi,\, s^k + \alpha_D \Delta s)\,.$$

Re-set $k \leftarrow k+1$ and return to (b).

5.4   Remarks on Interior-Point Methods

- The algorithm just described is almost exactly what is used in commercial interior-point method software.

- A typical interior-point code will solve a linear or quadratic optimization problem in 25-80 iterations, regardless of the dimension of the problem.

- These days, interior-point methods have been extended to allow for the solution of a very large class of convex nonlinear optimization problems.

6   Exercises on Interior-Point Methods for LP

1. (Interior Point Method) Verify that the formulas given in (11) indeed solve the equation system (10).

2. (Interior Point Method) Prove Proposition 6.1 of the notes on Newton's method for a logarithmic barrier algorithm for linear programming.

3. (Interior Point Method) Create and implement your own interior-point algorithm to solve the following linear program:

LP:
$$\text{minimize} \quad c^T x$$

$$\text{s.t.} \quad Ax = b$$
$$x \geq 0,$$

where
$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 100 \\ 50 \end{pmatrix},$$

and
$$c^T = \begin{pmatrix} -9 & -10 & 0 & 0 \end{pmatrix}.$$

In running your algorithm, you will need to specify the starting point (x_0, π^0, s^0), the starting value of the barrier parameter θ^0, and the value of the fractional decrease parameter α. Use the following values and make eight runs of the algorithm as follows:

- Starting Points:
  - $(x^0, \pi^0, s^0) = ((1, 1, 1, 1), (0, 0), (1, 1, 1, 1))$
  - $(x^0, \pi^0, s^0) = ((100, 100, 100, 100), (0, 0), (100, 100, 100, 100))$
- Values of θ^0:
  - $\theta^0 = 1000.0$
  - $\theta^0 = 10.0$
- Values of α:
  - $\alpha = 1 - \frac{1}{10\sqrt{n}}$
  - $\alpha = \frac{1}{10}$

4. (Interior Point Method) Suppose that $\bar{x} > 0$ is a feasible solution of the following logarithmic barrier problem:

$P(\theta)$:

$$\text{minimize} \quad c^T x - \theta \sum_{j=1}^{n} \ln(x_j)$$

$$\text{s.t.} \quad Ax = b$$
$$x \geq 0,$$

and that we wish to test if $\bar{x}$ is a β-approximate solution to $P(\theta)$ for a given value of θ. The conditions for $\bar{x}$ to be a β-approximate solution to $P(\theta)$ are that there exist values (π, s) for which:

$$\begin{cases} A\bar{x} = b, \bar{x} > 0 \\\\ A^T \pi + s = c \\\\ \|\frac{1}{\theta}\bar{X}s - e\| \leq \beta. \end{cases} \tag{17}$$

Construct an analytic test for whether or not such (π, s) exist by solving an associated equality-constrained quadratic program (which has a closed form solution). HINT: Think of trying to minimize $\|\frac{1}{\theta}\bar{X}s - e\|$ over appropriate values of (π, s).

5. Consider the logarithmic barrier problem:

$$P(\theta): \quad \min \quad c^T x - \theta \sum_{j=1}^{n} \ln x_j$$
$$\text{s.t.} \quad Ax = b$$

$$x > 0.$$

Suppose $\bar{x}$ is a feasible point, i.e., $\bar{x} > 0$ and $A\bar{x} = b$. Let $\bar{X} = \text{diag}(\bar{x})$.

a. Construct the (primal) Newton direction d at $\bar{x}$. Show that d can be written as:

$$d = \bar{X}(I - \bar{X}A^T(A\bar{X}^2 A^T)^{-1}A\bar{X})\left(e - \frac{1}{\theta}\bar{X}c\right).$$

b. Show in the above that $P = (I - \bar{X}A^T(A\bar{X}^2 A^T)^{-1}A\bar{X})$ is a projection matrix.

c. Suppose $A^T \bar{\pi} + \bar{s} = c$ and $\bar{s} > 0$ for some $\bar{\pi}, \bar{s}$. Show that d can also be written as:

$$d = \bar{X}(I - \bar{X}A^T(A\bar{X}^2 A^T)^{-1} A\bar{X}) \left(e - \frac{1}{\theta} \bar{X}\bar{s} \right).$$

d. Show that if there exists $(\bar{\pi}, \bar{s})$ with $A^T\bar{\pi} + \bar{s} = c$ and $\bar{s} > 0$, and $\|e - \frac{1}{\theta}\bar{X}\bar{s}\| \leq \frac{1}{4}$, then $\|\bar{X}^{-1}d\| \leq \frac{1}{4}$, and then the Newton iterate satisfies $\bar{x} + d > 0$.