Interior Point Methods

Alan Parks

This work was begun in April 05, revised beginning January 2010 and continuing into September 2010.

1 The Barrier Problems

We consider the LP: minimize $C \cdot X$ such that $A \cdot X = B$ and $X \geq \mathbb{O}$, where C is $1 \times n$ and A is $m \times n$ and B is $m \times 1$. Assume that the set of $X \in \mathbb{R}^n$ such that $A \cdot X = B$ is bounded.

An *interior point* is a *positive* vector that is feasible for the LP. We assume the existence of an interior point Y.

For a given positive number t, and for $X > \mathbb{O}$, define[1]

$$\beta_t(X) = C \cdot X - t \cdot \sum_{j=1}^{n} \ln(X_j)$$

Here is the associated *t-barrier problem.*

$$\text{Minimize} \quad \beta_t(X) \quad \text{such that} \quad A \cdot X = B \quad \text{and} \quad X > \mathbb{O}$$

In the barrier problem, t is constant.[2]

To see that the barrier problem always has a solution, we need the following.

---

[1] Regarding the logarithmic part of β_t, we think that the only properties needed are, 1) to go to $-\infty$ if some coordinate of X goes to 0, and 2) to be bounded above on the domain of the problem.

[2] Some sources use $1/t$ where we have used t. Obviously, it doesn't matter.

Lemma 1. *Given A, B, C as above, let ν be an upper bound on $|X|$ for all $X \in \mathbb{R}^n$ such that $A \cdot X = B$ and $X > \mathbb{O}$. Let $\mu \in \mathbb{R}$ with*

$$0 < \mu < \exp\Big[-\frac{1}{t} \cdot \beta_t(Y) - \frac{1}{t} \cdot \nu \cdot |C| - (n-1) \cdot \ln(\nu) \Big] \tag{1}$$

Let $X \in \mathbb{R}^n$ with $A \cdot X = B$ and $X \geq \mathbb{O}$. Suppose that $X_j < \mu$ for some j. Then $\beta_t(X) > \beta_t(Y)$.

Proof. Direct calculation.

$$\begin{aligned} \beta_t(X) &> -|C| \cdot |X| - t \cdot \ln(\mu) - t \cdot \sum_{k \neq j} \ln(X_k) \\ &\geq -|C| \cdot \nu - t \cdot (n-1) \cdot \ln(\nu) - t \cdot \ln(\mu) \\ &\geq -|C| \cdot \nu - -t \cdot (n-1) \cdot \ln(\nu) \\ &\qquad - t \cdot \left[-\frac{1}{t} \cdot \beta_t(Y) - \frac{1}{t} \cdot \nu \cdot |C| - (n-1) \cdot \ln(\nu) \right] \\ &= \beta_t(Y) \end{aligned}$$

as claimed. □

Proposition 2. *The t-barrier problem has a solution for each $t > 0$.*

Proof. Choose μ as in Lemma 1, and also be sure that $\mu \leq Y[k]$ for all k.

Consider the problem: minimize $\beta_t(X)$ such that $A \cdot X = B$ and $X[k] \geq \mu$ for all k. Notice that the condition on μ shows that Y is feasible for this problem. The domain of the problem is non-empty, closed, and bounded, and so the problem has a solution Z, and then $\beta_t(Z) \leq \beta_t(Y)$. If $A \cdot X = B$ and $X > \mathbb{O}$ but X is not in the domain, then Lemma 1 shows that $\beta_t(X) > \beta_t(Y) \geq \beta_t(Z)$. Therefore, Z is a solution to the barrier problem. □

Denote by X_t a solution to the t-barrier problem.

Proposition 3. *Let $0 < s < t$. Then $C \cdot X_t \geq C \cdot X_s$, and*

$$\sum_{k=1}^{n} \ln(X_s[k]) \leq \sum_{k=1}^{n} \ln(X_t[k])$$

Proof. The vector X_t is feasible for the s-barrier problem, and so

$$C \cdot X_s - s \cdot \sum_{k=1}^{n} \ln(X_s[k]) \le C \cdot X_t - s \cdot \sum_{k=1}^{n} \ln(X_t[k])$$

Similarly, X_s is feasible for the t-barrier problem.

$$C \cdot X_t - t \cdot \sum_{k=1}^{n} \ln(X_t[k]) \le C \cdot X_s - t \cdot \sum_{k=1}^{n} \ln(X_s[k])$$

Putting this together, we obtain

$$\frac{1}{t} \cdot C \cdot (X_t - X_s) \le \sum_{k=1}^{n} \ln(X_t[k]/X_s[k]) \le \frac{1}{s} \cdot C \cdot (X_t - X_s)$$

and so

$$\begin{aligned} \frac{1}{t} \cdot C \cdot (X_t - X_s) &\le \frac{1}{s} \cdot C \cdot (X_t - X_s) \\ 0 &\le \left[\frac{1}{s} - \frac{1}{t}\right] \cdot C \cdot (X_t - X_s) \end{aligned}$$

Since $s < t$, we see that $1/t < 1/s$, and so $C \cdot X_s \le C \cdot X_t$; this is the first conclusion of this proposition.

Since $C \cdot (X_t - X_s) \ge 0$, we have $\frac{1}{t} \cdot C \cdot (X_t - X_s) \ge 0$. It follows that

$$0 \le \sum_{k=1}^{n} \ln(X_t[k]/X_s[k])$$

so that the second conclusion holds. □

The vectors X_t are bounded, and so they have a limit point Z. There is therefore a set S of positive real numbers, having 0 as a limit point, and such that $X_t \to Z$ for $t \to 0$ with $t \in S$.

Proposition 4. *The vector Z is a solution to the LP.*

Proof. Because $A \cdot X_t = B$ and $X_t \ge \mathbb{O}$ for each t, we have $A \cdot Z = B$ and $Z \ge \mathbb{O}$, so that Z is feasible for the LP. Proposition 3 shows that $C \cdot X_t$ decreases as $t \to 0$, and so $C \cdot Z \le C \cdot X_t$ for all $t \in S$.

Suppose that V is an interior point; we claim that $C \cdot Z \leq C \cdot V$. Indeed, suppose that $C \cdot V < C \cdot Z$. Since $\beta_t(V) \to C \cdot V$ as $t \to 0$, there is $t_0 > 0$ such that if $0 \leq t \leq t_0$, then

$$\beta_t(V) < C \cdot Z$$

Since V is an interior point, we have $\beta_t(X_t) \leq \beta_t(V)$ for all $t > 0$. Restricting to $t \in S$, we have

$$\beta_t(X_t) \leq \beta_t(V) < C \cdot Z \leq C \cdot X_t$$

This shows that

$$0 \leq C \cdot X_t - \beta_t(X_t) = t \cdot \sum_{k=1}^{n} \ln(X_t[k])$$

Since each $X_t[k]$ is bounded above by ν (from Lemma 1), we see that the sum of $\ln(X_t[k])$ is bounded above. It follows that

$$\lim_{t \to 0} \left[t \cdot \sum_{k=1}^{n} \ln(X_t[k]) \right] = 0$$

Since also $C \cdot X_t \to C \cdot Z$ as $t \to 0$ (with $t \in S$), we see that $\beta_t(X_t) \to C \cdot Z$ as $t \to 0$. We already had $\beta_t(V) \to C \cdot C$ as $t \to 0$. Since $\beta_t(X_t) \leq \beta_t(V)$, we see that $C \cdot Z \leq C \cdot V$, a contradiction. The claim holds.

Now suppose that V is a solution to the LP, yet $C \cdot V < C \cdot Z$. Recall the interior point Y, and observe that $s \cdot Y + (1 - s) \cdot V$ is an interior point when $0 < s \leq 1$. We can pick s small enough so that

$$C \cdot (s \cdot Y + (1 - s) \cdot V) < C \cdot Z$$

and this contradicts the claim. Therefore, Z solves the LP. □

2 The Barrier Equations

We now describe how to approximate a solution to the t-barrier problem in the case that the $m \times n$ matrix A has rank m. Then a solution will necessarily come from a solution to the Kuhn-Tucker conditions. Thus, for each $t > 0$, there will be a $1 \times m$ matrix L_t such that

$$C - t \cdot \begin{pmatrix} \frac{1}{X_t[1]} & \frac{1}{X_t[2]} & \cdots & \frac{1}{X_t[n]} \end{pmatrix} = L_t \cdot A \tag{2}$$

There are no sign conditions on L_t. A generic interior point method tries to solve (2) for a succession of decreasing values of t.

Suppose that X_t is a solution to (2). Define

$$S_t[i] = t \cdot \frac{1}{X_t[i]} \quad \text{for} \quad 1 \leq i \leq n$$

and then (2) says that $C = S_t + L_t \cdot A$, where $S_t > \mathbb{O}$. In other words, (S_t, L_t) is a feasible vector for the LP that is dual to the original (primal) LP.

We can summarize the t-barrier problem by saying that it looks for X, L, S such that

$$A \cdot X = B \tag{3}$$

$$L \cdot A + S = C \tag{4}$$

$$S[i] \cdot X[i] = t \quad \text{for each } i \tag{5}$$

$$X, S > \mathbb{O} \tag{6}$$

Matrix dimensions: X is $n \times 1$ and L is $1 \times m$ and S is $1 \times n$.

3 Our Canonical Form

3.1 Set-up

We establish notation that we will use for various numerical experiments.[3] We will be using a set-up equivalent to that of Wright's book [1], except that we will assume that the equation constraint of our LP is in reduced form – equivalently, we assume that our constraints are in inequality form.

Our Canonical LP: minimize $C \cdot X_1$ such that $A \cdot X_1 + X_2 = B$ and $X = (X_1, X_2) \geq \mathbb{O}$. Let A be $m \times n$, so that C is $1 \times n$. As in [1] we assume there is a feasible vector $Y = (Y_1, Y_2) > \mathbb{O}$ with $A \cdot Y_1 + Y_2 = B$, so that Y is an interior point in equation form.

Equation (4) takes the form of two equations, splitting $S = (S_1, S_2)$.

$$L \cdot A + S_1 = C$$
$$L + S_2 = \mathbb{O}$$

---

[3] This was changed on September 2. Previously we had been using free variables in (3) to enforce a solution to that equation, and to solve (5). We wrote (4) in terms of the free vector, and used Newton's Method to solve that equation, all the while maintaining (6).

where $S > \mathbb{O}$. We can drop L in favor of the single equation $S_1 = S_2 \cdot A + C$, where S_1 is $1 \times n$ and S_2 is $1 \times m$. We assume we are given $Z_2 > \mathbb{O}$ such that $Z_2 \cdot A + C > \mathbb{O}$, so that if we define $Z_1 = Z_2 \cdot A + C$, then $Z = (Z_1, Z_2)$ is positive. Thus, in the present case, equations (3)-(6) amount to these:

$$A \cdot X_1 + X_2 = B \tag{7}$$

$$S_1 - S_2 \cdot A = C \tag{8}$$

$$S[i] \cdot X[i] = t \quad \text{for all} \quad 1 \leq i \leq n + m \tag{9}$$

$$S, X > \mathbb{O} \tag{10}$$

We view the left sides as a function of X_1, S_2. We assume we have vectors Y_1, Z_2, playing the roles of X_1, S_2, respectively, that produce Y_2, Z_1 and satisfy (7) and (8) and (10). Thus, to solve the barrier problem, we need to obtain (9).

3.2 Newton's Method

For each positive integer k, let J_k be the $k \times 1$ matrix all of whose entries are 1.

We regard the left side of (9) as giving a function of X_1, S_2; as such it maps $\mathbb{R}^{n+m}$ to itself, and so Newton's Method raises its hand. Wright (p.7 of [1]) introduces the *centering parameter* μ; we will use the Roman name M. In our present context,

$$(m + n) \cdot M = \sum_{j=1}^{n+m} S[i] \cdot X[i] \tag{11}$$

For the term t in the barrier equation, we substitute $s \cdot M$, where $s \in [0, 1]$. Since M is determined by X_1, S_2, the condition $s < 1$ will move us toward a solution to the LP – that solution occurs as $t \longrightarrow 0$. The equations (9) and are replaced by

$$S[i] \cdot X[i] = s \cdot M \quad \text{for} \quad 1 \leq i \leq n + m$$

Remembering that X_2, S_1 are functions of X_1, S_2, Newton's Method produces this iteration:

$$\begin{bmatrix} \delta(S_1) & \delta(X_1) \cdot A^T \\ -\delta(S_2) \cdot A & \delta(X_2) \end{bmatrix} \cdot \begin{bmatrix} \Delta X_1 \\ \Delta S_2 \end{bmatrix} = \begin{bmatrix} -\delta(S_1) \cdot X_1 + s \cdot M \cdot J_n \\ -\delta(X_2) \cdot S_2 + s \cdot M \cdot J_m \end{bmatrix} \tag{12}$$

Since S, X are positive, the coefficient matrix on the left is invertible.

To maintain positive vectors (to maintain an interior point), we scale Newton's deltas by a positive number p, so that for each variable W, we have $W + p \cdot \Delta W > \mathbb{O}$. (Here $W = X, S$.) So the actual iteration delta is $p \cdot \Delta W$.

3.3 The Capital Basic Vector

Suppose that Z is a basic solution (for X) to the primal LP of this section. Then Z is determined by a choice of m basic variables; say those variables are indexed by the set J of m elements of $\{1, 2, \ldots, n+m\}$. Since the coordinates of Z at the free variables are 0, we see that the values $Z[j]$ for $j \in J$ are the m largest coordinates of Z. Turning this around, let Z be an arbitrary feasible vector (for X) in the LP. Define the *capital* of Z to be the set J of the indices of the m largest coordinates of Z. (If there are equal coordinates, then the capital may not be well-defined; we will not worry about this.) Given the capital J, we can attempt to put the LP into a reduced form where the variables $X[j]$ with $j \in J$ are basic for the equation constraint. If this is possible, we obtain the *capital basic vector* Z' for X. We can observe whether the reduced form, including the reduced form of the objective equation, is in solution form. Since the barrier solutions approach a solution to the LP as $t \to 0$ (as $s \cdot M \to 0$), we suspect that some capital basic vector will be a solution to the LP.

3.4 Random Examples

We can generate a random $X > \mathbb{O}$ and a random A, and then B is defined. We can generate a random $S > \mathbb{O}$, and then C is defined. We might wish to have $X[i] \cdot S[i] = 1$ for all i, so that we begin with a solution to the 1-barrier problem.

References

[1] Stephen J. Wright, *Primal-Dual Interior-Point Methods*, Philadelphia: S.I.A.M. 1997.