Journal of Computational Mathematics, Vol.19, No.1, 2001, ??–??.

A SCALED CENTRAL PATH FOR LINEAR PROGRAMMING*[1)]

Ya-xiang Yuan
(*State Key Laboratory of Scientific and Engineering Computing,*
Institute of Computational Mathematics and Scientific/Engineering Computing,
Chinese Academy of Sciences, Beijing 100080, China)

Abstract

Interior point methods are very efficient methods for solving large scale linear programming problems. The central path plays a very important role in interior point methods. In this paper we propose a new central path, which scales the variables. Thus it has the advantage of forcing the path to have roughly the same distance from each active constraint boundary near the solution.

Key words: Central path, Interior point methods, Linear programming.

1. Introduction

Interior point methods are one of the most intensively studied topics in optimization. Thousands of publications have been appeared on interior point methods. A very good recent review is given by [4]. Interior point methods have very good theoretical properties including the nice polynomial complexity property. And more important is that numerous applications have shown that interior point methods are very efficient for solving large sparse linear programming problems. Interior point methods have been proved to be indispensable to semi-definite programming, another class of important optimization problems. Interior point methods have also been applied to nonlinear programming and nonlinear complementary problems. For examples of detailed discussions, please see ([1, 2, 3]).

Path following algorithms are a class of very important interior point methods for linear programming. Consider the following standard linear programming problem

$$\min c^T x \tag{1.1}$$

subject to

$$Ax = b, \qquad x \geq 0, \tag{1.2}$$

where $c \in \Re^n$, $b \in \Re^m$ and $A \in \Re^{m \times n}$. The dual problem for the above linear program can be written as

$$\max b^T y \tag{1.3}$$

subject to

$$A^T y + s = c, \qquad s \geq 0, \tag{1.4}$$

where $y \in \Re^m$ are the dual variables and $s \in \Re^n$ are the slack variables. If both the prime problem (1.1)-(1.2) and the dual problem (1.3)-(1.4) have feasible solutions, then both problems have optimal solutions. And, in this case, for any solution x^* of the primal problem and any solution (y^*, s^*) of the dual problem, we have that

$$c^T x^* = b^T y^*, \tag{1.5}$$

* Received November 15, 2000.
[1)]This work is partially supported by Chinese NNSF grants 19731010 and the Knowledge Innovation Program of Chinese Academy of Sciences.

(see, [5]). For any point (x, y, x) that satisfies (1.2) and (1.4), it follows that

$$c^T x - b^T y = (A^T y + s)^T x - (Ax)^T y = s^T x \geq 0. \tag{1.6}$$

Thus, a solution is obtained as long as the complementarity gap $s^T x$ is zero. Because both x and s are nonnegative, the condition $s^T x = 0$ is equivalent to $x_i s_i = 0$ for all $i = 1, ..., n$. Let $X = diag[x_1, x_2, ..., x_n]$, relation $s^T x = 0$ can be expressed as $Xs = 0$. Thus, we can write the optimal conditions in the following form

$$\begin{aligned} Ax &= b && (1.7)\\ A^T y + s &= c && (1.8)\\ Xs &= 0 && (1.9)\\ (x, s) &\geq 0. && (1.10) \end{aligned}$$

Define the set

$$\mathcal{F} = \{(x, y, s)\ Ax = b, A^T y + s = c, x \geq 0, s \geq 0\}, \tag{1.11}$$

which is the direct product of the primal feasible set and the dual feasible set. Interior point methods generate iterate point in the interior of the region $\mathcal{F}$, that is

$$int(\mathcal{F}) = \{(x, y, s)\ Ax = b, A^T y + s = c, x > 0, s > 0\}. \tag{1.12}$$

The central path is defined by the following system

$$\begin{aligned} Ax &= b && (1.13)\\ A^T y + s &= c && (1.14)\\ Xs &= \mu e && (1.15)\\ (x, s) &> 0, && (1.16) \end{aligned}$$

where e is a vector whose elements are all 1, and $\mu > 0$ is a parameter. It is easy to see that system (1.13)-(1.16) is a perturbation of the optimal condition (1.7)-(1.10). Let $(x(\mu), y(\mu), s(\mu))$ be on the central path, it can be shown that $x(\mu)$ is a solution of the penalized problem

$$\min\ c^T x\ -\ \mu \sum_{i=1}^{n} \log(x_i) \tag{1.17}$$

subject to

$$Ax = b. \tag{1.18}$$

Many interior point methods use the central path. Some algorithms explicitly use the central path as they force the interate points to follow the central path. Even for many algorithms that do not use the central path directly in the algorithm statements, the central path is used for convergence analyses (see, [5]).

Because of the importance of the central path in the designs and analyses of interior point methods, we study the center path. From the views of complementary conditions, we propose a new central path. We believe that this path can also be used to construct new interior point methods.

A new central path is derived in the next section, and in Section 3 we proposed two ways to compute search directions based on the new central path and in Section 4 we give a brief discussion on how our ideas can be further extended.

2. A Scaled Central Path

The condition (1.15) of the central path requires that all elements of Xs are the same when they approach zero. This can be viewed as that the complementarity conditions

$$x_i s_i = 0, \qquad (i = 1, ..., n) \tag{2.1}$$

are replaced or approximated by the relation

$$x_i s_i = \mu, \qquad (i = 1, ..., n). \tag{2.2}$$

This is reasonable because if some elements of $x_i s_i$ approach zero much faster than the other elements, the point (x, y, s) is much closer to the boundary of feasible set of (1.13), (1.10) and (1.3) than to the solution (x^*, y^*, s^*). Roughly speaking, condition (2.2) prevents the point (x, y, s) from being too close to any particular active boundary that has the form $x_i = 0$ or $s_i = 0$.

However, a second look on the central path condition (2.2) reveals that there are rooms to make an improvement. Suppose at the solution point (x^*, y^*, s^*) the strictly complementary conditions hold, namely

$$\tau_i^* = x_i^* + s_i^* > 0, \qquad (i = 1, ..., n). \tag{2.3}$$

When both the primal problem and the dual problem are feasible, such solutions always exist (see, [5]). Assume that J^* is the subset of $\{1, 2, ..., n\}$ such that

$$x_i^* = 0, \quad i \in J^* \qquad \text{and} \qquad s_i^* = 0, \quad i \notin J^*. \tag{2.4}$$

Thus, when a point (x, y, s) on the central path is close to the solution (x^*, y^*, s^*), the distance from the iterate point (x, y, s) to each active boundary near the solution is

$$x_i = \frac{\mu}{s_i} \approx \frac{\mu}{s_i^*}, \quad (i \in J^*), \tag{2.5}$$

and

$$s_i = \frac{\mu}{x_i} \approx \frac{\mu}{x_i^*}, \quad (i \notin J^*). \tag{2.6}$$

From (2.3), we can see that the distances from the central path to the active boundaries are roughly $\mu/\tau_i^*(i = 1, ..., n)$. These distances will be different because generally the numbers τ_i^* need not to be the same. Hence, in this sense, we can say that the central path is not exactly central as the distances to the active boundaries are not the same.

If we require that

$$x_i = \mu, \quad i \in J^* \qquad \text{and} \qquad s_i = \mu, \quad i \notin J^*, \tag{2.7}$$

when the point (x, y, s) is close to the solution, the distances from this point to all active boundaries are the same. But, normally the set J^* is not known before the problem is solved. Strict complementarity conditions (2.3) indicate that near the solution relations (2.7) are equivalent to

$$\min[x_i, s_i] = \mu, \quad (i = 1, ..., n). \tag{2.8}$$

The path that satisfies (1.13), (1.14), (1.16) and (2.8) can be regarded as a "strictly central" path of the feasible region (1.12) near the solution (x^*, y^*, s^*) in the sense that it has the same distance to all active boundaries. Unfortunately the function $\min[x_i, s_i]$ is a nonsmooth function. Therefore we use the following approximation

$$\frac{x_i s_i}{x_i + s_i} \approx \min[x_i, s_i]. \tag{2.9}$$

This formula is a very good approximation if $\min[x_i, s_i] << \max[x_i, s_i]$, which is exactly the case when (x, s) is close to a solution (x^*, s^*) at which (2.3) holds.

Thus, we can define a new central path by the following system

$$\begin{aligned} Ax &= b && (2.10)\\ A^T y + s &= c && (2.11)\\ Xs &= \mu(x+s) && (2.12)\\ (x, s) &> 0. && (2.13) \end{aligned}$$

We call this central path the *scaled central path.*

Similar to the central path (1.13)-(1.16), the scaled central path also has a *log* penalty function property.

Theorem 2.1 *Let $(x(\mu), y(\mu), s(\mu))$ be a point in the scaled central path (2.10)-(2.13), then $x(\mu)$ is a minimum of*

$$\min c^T x - \mu \sum_{i=1}^{n} [x_i + \mu \log(x_i - \mu)] \tag{2.14}$$

subject to

$$Ax = b. \tag{2.15}$$

Proof Because both $x(\mu)$ and $s(\mu)$ are positive, it follows from (2.12) that

$$x_i(\mu) > \mu. \tag{2.16}$$

Define

$$h(x) = c^T x - \mu \sum_{i=1}^{n} [x_i + \mu \log(x_i - \mu)]. \tag{2.17}$$

From (2.12), we have that

$$s(\mu) = \mu(X(\mu) - \mu I)^{-1} x(\mu), \tag{2.18}$$

where $X(\mu) = diag(x_1(\mu), ..., x_n(\mu))$. Thus, we have that

$$s(\mu) = c - \frac{d}{dx} h(x(\mu)). \tag{2.19}$$

The above relation and (2.11) indicates that

$$\frac{d}{dx} h(x(\mu)) = A^T y. \tag{2.20}$$

Equations (2.20) and (2.10) show that $x(\mu)$ is a Kuhn-Tucker point of problem (2.14)-(2.15). It is easy to see that function $h(x)$ is a convex function for all x that satisfy $x_i > \mu$. Therefore $x(\mu)$ is the unique minimum of problem (2.14)-(2.15) on the region $\{x | x_i > \mu, i = 1, ..., n.\}$. $\square$

The object function (2.14) can be viewed as the sum of a slightly perturbation of the original objective function $(c - \mu e)^T x$ and a shifted penalty function $\sum \log(x_i - \mu)$. However, it is not clear what are the advantages of using the penalty parameter to perturbate the objective function and to shift the *log* penalty functions.

It is interesting to study the theoretical properties of the scaled central path, and numerical behaviour of the interior point methods based on the scaled central path. In the following, we demonstrate that the scaled central path can be used to obtain search directions for solving linear programming problems.

3. Search Directions

Interior point methods following the centeral path generate points on or near the central path (1.13)-(1.16) with μ being reduced every iteration. For example, suppose at the k-th iteration, (x_k, s_k, y_k) is on the central path (1.13)-(1.16), we hope that the next iterate is still on the central path with μ being replaced by $\gamma\mu$, where $\gamma \in (0,1)$ is a constant. Thus search directions can be obtained by applying Newton's method to the following system:

$$\begin{aligned} Ax &= b & (3.1)\\ A^T y + s &= c & (3.2)\\ Xs &= \gamma\mu. & (3.3) \end{aligned}$$

For more details, please see [5].

Based on this approach, we can show that the scaled central path (2.10)-(2.13) can be used to compute search directions. Assume the current iterate point (x, y, s) is in the set (1.12). Let the search direction be (d_x, d_y, d_s). Directly applying the Newton's method to the scaled central path (2.10)-(2.13) (μ replaced by $\gamma\mu$), we obtain:

$$\begin{aligned} Ad_x &= 0 & (3.4)\\ A^T d_y + d_s &= 0 & (3.5)\\ (X - \gamma\mu I)d_s + (S - \gamma\mu I)d_x &= \gamma\mu(x + s) - Xs, & (3.6) \end{aligned}$$

where $S = diag[s_1, s_2, ..., s_n]$. Here

$$\mu = x^T s/(\|x\|_1 + \|s\|_1), \qquad (3.7)$$

and γ is a positive number in $(0,1)$. We can rewrite (2.12) (μ replaced by $\gamma\mu$) as

$$X(X + S)^{-1}s = \gamma\mu e. \qquad (3.8)$$

Replacing (2.12) by (3.8) and applying Newton's method again, we obtain

$$\begin{aligned} Ad_x &= 0 & (3.9)\\ A^T d_y + d_s &= 0 & (3.10)\\ S^2 d_x + X^2 d_s &= \gamma\mu(X + S)(x + s) - XS(x + s). & (3.11) \end{aligned}$$

Here

$$\mu = x^T (X + S)^{-1} s/n \qquad (3.12)$$

and γ, as above, is a positive number in $(0,1)$.

4. Discussions

We can generalize (2.14)-(2.15) to the following penalty function problem:

$$\min(c - \mu_1 e)^T x - \mu \sum_{i=1}^{n} \log(x_i - \mu_2) \qquad (4.1)$$

subject to

$$Ax = b. \qquad (4.2)$$

The scaled central path corresponds to the case when $\mu_1 = \mu_2 = \sqrt{\mu}$. And the standard central path corresponds to the case when $\mu_1 = \mu_2 = 0$. It is obvious that new central paths can be obtained if we choose different μ_1 and μ_2.

Now let us have a closer look at the technique that motivates the scaled central path. The key idea is to approximate the strict central path conditions (2.8) by smooth equations. This is a problem of choosing a smooth function $\phi(x,s)$ defined in $\Re^2_+$ such that

$$\phi(x,s) \approx \min(x,s), \qquad \forall x>0, s>0. \tag{4.3}$$

If the above ralation is true, function $\psi(x,s)$ defined by $\psi(x.s) = xs/\phi(x,s)$ is an approximation of $\max(x,s)$. Since $\max(x,s) = [\min(x^{-1},s^{-1})]^{-1}$, if

$$\psi(x,s) = [\phi(x^{-1},s^{-1})]^{-1}, \tag{4.4}$$

we call $\phi(x,s)$ is a self-dual approximation to $\min(x,s)$. Relation (4.4) indicates that approximations for $\min(x,s)$ and $\max(x,s)$ are symmetric. (4.4) is equivalent to

$$\phi(x^{-1},s^{-1}) = \phi(x,s)/xs. \tag{4.5}$$

The function

$$\phi_1(x,s) = \frac{xs}{x+s} \tag{4.6}$$

that we use to obtain the scaled central path satisfies condition (4.5). We can easily see that $\phi_1(x,s)$ always approximates $\min(x,s)$ from below. If $x \approx s$, $\phi(x,s)$ is not a good approximation of $\min(x,s)$. In the extreme case when $s=x$, $\phi(x,s) = \min(x,s)/2$. An interesting question is what is the best smooth approximation to the nonsmooth function $\min(x,s)$. In the following we give another approximation formula:

$$\phi_2(x,s) = \frac{xs}{x+s} + \frac{4x^2s^2}{(x+s)^3}, \tag{4.7}$$

which also satisfy the dual condition (4.5). One nice property of the the above formula is that it satisfies

$$\min(x,s) \le \phi_2(x,s) \le \frac{(x+s)}{2}. \tag{4.8}$$

Hence it seems that (4.7) is a better approximation to $\min(x,s)$ than (4.6). But, unfortunately, (4.7) is more complicated than (4.6). Therefore it is not obvious whether efficient methods can be constructed based on the central path defined by (4.7), (1.13), (1.14) and (1.16).

We have proposed a new central path for linear programming. This new central path has the property that the distances from all active boundaries are nearly the same when the path is close to the solution. We believe that this central path can be used to construct new efficient interior point methods and to analyze theoretical properties of many interior point methods. Our analyses also indicate that our ideas can be extended to define different central paths.

References

[1] F. Alizadeh, J.-P.A. Haeberly, M.L. Overton, A new primal-dual interior point method for semidefinite programming, in: J.G. Lewis, ed., *Proceedings of the Fifth SIAM Conference on Applied Linear Algebra*, (SIAM, Philadelphia, 1994), 113-117.

[2] T.F. Coleman, Y. Li, A trust region and affine scaling interior point method for nonconvex minimization with linear inequality constraints, *Math. Program.,* **88** (2000), 1-31.

[3] J.Y. Han, Y.B. Zhao, Two interior-point methods for nonlinear $P_*(\tau)$-complementarity problems, *J. Optim. Theory Appl.*, (1999), 659–679.

[4] S.J. Wright, Recent developments in interior-point methods, in: M.J.D. Powell and S. Scholtes, eds., *System Modelling and Optimization*, (Kluwer, Boston, 2000), 311–333.

[5] Y. Ye, Interior Point Algorithms: Theory and Analysis, John Wiley and Sons, Chichester, 1997.