

90-B-4

AN ACTIVE-SET STRATEGY
IN INTERIOR POINT METHOD
FOR LINEAR PROGRAMMING

Kaoru TONE

Revised

January 4, 1991

AN ACTIVE-SET STRATEGY IN INTERIOR POINT METHOD FOR LINEAR PROGRAMMING*

Kaoru TONE
Graduate School of Policy Science
Saitama University
Urawa, Saitama 338, Japan.

January 4, 1991

Abstract

We will present a potential reduction method for linear programming where only the constraints with relatively small dual slacks — active constraints— will be taken into account to form the ellipsoid constraint at each iteration of the process. The algorithm converges to the optimal feasible solution in $O(\sqrt{n}L)$ iterations with the same polynomial bound with the full constraints case, where n is the number of variables and L is the data length. If a small portion of the constraints is active near the optimal solution, the computational cost to find the next direction of movement in one iteration will be fairly reduced by the proposed strategy. As a special case of this strategy, we will show that the interior point method can be managed by the basis factorization techniques of the simplex method coupled with a sequence of rank-one changes to matrices.

Key words: Linear programming, interior point method, active set strategy.

Abbreviated title: Active-set LP algorithm

*This research was partially done in June 1990 while the author was visiting Department of Mathematics, University of Pisa.

1 Introduction

Since the epoch-making breakthrough by Karmarkar [6], the interior point methods for linear programming have been studied in many aspects. Karmarkar introduced the potential function for linear programming for the first time. Since then many researchers have studied and extended it with big theoretical successes. (See Gonzaga [4], Ye [10], Freund [3], Todd and Ye [8], Kojima, Mizuno and Yoshise [7] among others.) Ye [10] has developed an $O(\sqrt{n}L)$ iteration potential reduction algorithm based on the primal-dual potential function. (Gonzaga [5] and Freund [3] have presented similar results.)

The purpose of this paper is to introduce an active-set strategy in Ye's dual potential reduction method. Usually, a linear program in the inequality constraints form has many redundant constraints in the optimal solution. If we deal with all constraints, we are forced to solve simultaneous linear equations related with the entire coefficient matrix of the problem. The proposed active-set strategy will allow to reduce the cost fairly, still keeping the same order of polynomial convergence property. In Section 2, we will define the problem and the potential functions along with some preliminary lemmas. We mainly concern with the dual form of linear programming. In Section 3, an active set strategy for finding the direction of movement in the scaled dual space will be introduced. Also, we will discuss on some properties related with the expansion of the active set. In Section 4, we will introduce the primal variable. In Section 5, an algorithm based on the active set strategy will be presented. In Section 6, we will analyze the convergence property of the algorithm and show that it terminates in $O(\sqrt{n}L)$ iterations. In Section 7, we will observe the case when the active set constitutes a basis and show that the interior point method can be managed by the basis factorization techniques of the simplex method coupled with a sequence of rank-one changes to matrices.

2 Problem and Potential Functions

We are concerned with the linear program whose primal form is

$$\langle P \rangle \text{ minimize } \bar{z} = c^T x \tag{2.1}$$

subject to $x \in \{x \in R^n : Ax = b, x \geq 0\}$,

where $c \in R^n$, $A \in R^{m \times n}$, and $b \in R^m$ are given.

Its dual form is

$$\langle D \rangle \text{ maximize } \underline{z} = y^T b \quad (2.2)$$

$$\text{subject to } y \in \{y \in R^m : c - A^T y \geq 0\}.$$

We denote the dual slack variables by

$$s = c - A^T y \geq 0. \quad (2.3)$$

As far as notations are concerned, the superscript T denotes the transpose, e the vector of all ones and A_j the j -th column vector of A . The upper case letter (X) designates the diagonal matrix of the vector (x) in lower case. The optimum value of \bar{z} (and hence \underline{z}) is denoted by z^* .

For $\langle P \rangle$ and $\langle D \rangle$, we assume that

1. the relative interiors of the feasible regions of $\langle P \rangle$ and $\langle D \rangle$ are nonempty and we have an interior feasible x^0 and y^0 for $\langle P \rangle$ and $\langle D \rangle$ such that

$$Ax^0 = b, \quad x^0 > 0 \quad (2.4)$$

and

$$s^0 = c - A^T y^0 > 0, \quad (2.5)$$

2. A has full row rank and has no zero columns,
3. the objective functions of $\langle P \rangle$ and $\langle D \rangle$ are not a constant on the feasible regions and
4. each A_j ($j = 1, \dots, n$) has unit 2-norm.

$$\|A_j\| = 1. \quad (j = 1, \dots, n) \quad (2.6)$$

Associated with the programs, we consider three kinds of potential function:

the *primal potential function* for an interior feasible x :

$$\phi_P(x, \underline{z}) = \rho \ln(c^T x - \underline{z}) - \sum_{j=1}^n \ln(x_j), \quad (2.7)$$

the *dual potential function* for an interior feasible (y, s) :

$$\phi_D(y, \bar{z}) = \rho \ln(\bar{z} - b^T y) - \sum_{j=1}^n \ln(s_j), \quad (2.8)$$

and

the *primal-dual potential function* for an interior feasible pair (x, y, s) :

$$\phi_{PD}(x, s) = \rho \ln(x^T s) - \sum_{j=1}^n \ln(x_j s_j), \quad (2.9)$$

where \underline{z} (\bar{z}) is a lower (upper) bound to z^* and ρ is a positive parameter.

We have the relations:

$$\phi_{PD}(x, s) = \phi_P(x, \underline{z}) - \sum_{j=1}^n \ln(s_j) \quad (2.10)$$

and

$$\phi_{PD}(x, s) = \phi_D(y, \bar{z}) - \sum_{j=1}^n \ln(x_j), \quad (2.11)$$

for $\underline{z} = b^T y$ and $\bar{z} = c^T x$.

The gradient vectors of ϕ_P (with x) and ϕ_D (with y) are

$$\nabla \phi_P(x, \underline{z}) = \frac{\rho}{c^T x - \underline{z}} c - X^{-1} e \quad (2.12)$$

and

$$\nabla \phi_D(y, \bar{z}) = -\frac{\rho}{\bar{z} - b^T y} b + AS^{-1} e. \quad (2.13)$$

The following lemma is due to Ye [10].

Lemma 1 (Ye) *For any two points*
 $s^0 = c - A^T y^0 > 0$, $s^1 = c - A^T y^1 > 0$ *and*

$$\|(S^0)^{-1}(s^1 - s^0)\|_\infty < 1, \quad (2.14)$$

we have

$$\begin{aligned} \phi_D(y^1, \bar{z}) - \phi_D(y^0, \bar{z}) &\leq \nabla \phi_D^T(y^0, \bar{z})(y^1 - y^0) \\ &+ \frac{\|(S^0)^{-1}(s^1 - s^0)\|^2}{2(1 - \|(S^0)^{-1}(s^1 - s^0)\|_\infty)}. \end{aligned} \quad (2.15)$$

3 Active Set Strategy

From the assumption above, we have an interior feasible pair (x^0, y^0, s^0) which satisfies

$$Ax^0 = b^0, \quad x^0 > 0, \quad A^T y^0 + s^0 = c, \quad s^0 > 0$$

with the primal objective value $\bar{z}^0 = c^T x^0$ and the dual objective value $\underline{z}^0 = b^T y^0$.

We partition the index set $J = \{1, 2, \dots, n\}$ into β and $\bar{\beta}$ in the following way:

1. Let us reorder the index set J in the ascending order of s_j^0 :

$$s_1^0 \leq s_2^0 \leq \dots \leq s_n^0. \quad (3.1)$$

2. Take A_1, A_2, \dots in this order into the set β until they span R^m . Let the last index be k and $\mu = s_k^0$. Then we have

$$\beta = \{j | s_j^0 \leq \mu\} \quad (3.2)$$

and

$$\bar{\beta} = J - \beta. \quad (3.3)$$

We will call the set β thus obtained as *active*. Notice that $s_j^0 > \mu$ ($j \in \bar{\beta}$) and $\bar{\beta}$ may be empty. We will use the suffix β ($\bar{\beta}$) to denote the partition of a matrix or a vector in accordance with the set β ($\bar{\beta}$).

Now, we minimize the linearized dual potential function subject to the ellipsoid constraint formed by the active set β .

$$\text{minimize } \nabla \phi_D^T(y^0, \bar{z}^0)(y - y^0) \quad (3.4)$$

$$\text{subject to } \|(S_\beta^0)^{-1} A_\beta^T (y - y^0)\| \leq \gamma. \quad (3.5)$$

The solution $y(\gamma)$ of this problem is given by

$$y(\gamma) = y^0 - \gamma d_\beta^0, \quad (3.6)$$

where the direction vector d_β^0 is defined as

$$d_\beta^0 = \frac{\Gamma_\beta^{-1} \nabla \phi_D(y^0, \bar{z}^0)}{\sqrt{\nabla \phi_D^T(y^0, \bar{z}^0) \Gamma_\beta^{-1} \nabla \phi_D(y^0, \bar{z}^0)}}, \quad (3.7)$$

with

$$\Gamma_\beta = A_\beta (S_\beta^0)^{-2} A_\beta^T. \quad (3.8)$$

The slack variable $s(\gamma)$ is given by

$$s(\gamma) = s^0 - A^T (y(\gamma) - y^0). \quad (3.9)$$

It is easy to see that

$$\|(S^0)^{-1} A^T d_\beta^0\|^2 = \|(S_\beta^0)^{-1} A_\beta^T d_\beta^0\|^2 + \|(S_{\bar{\beta}}^0)^{-1} A_{\bar{\beta}}^T d_\beta^0\|^2. \quad (3.10)$$

Lemma 2 $\|(S_\beta^0)^{-1} A_\beta^T d_\beta^0\| / \|(S^0)^{-1} A^T d_\beta^0\|$ is bounded below by a positive value.

Proof. Let the minimum eigenvalue of $A_\beta(S_\beta^0)^{-2}A_\beta^T$ for the active set β be λ_β which is positive since A_β is full row rank. Thus,

$$\|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\| \geq \sqrt{\lambda_\beta} \|d_\beta^0\|. \quad (3.11)$$

On the other hand, let the maximum eigenvalue of $A_{\bar{\beta}}(S_{\bar{\beta}}^0)^{-2}A_{\bar{\beta}}^T$ for $\bar{\beta}$ be $\bar{\lambda}_{\bar{\beta}}$ which is positive if $\bar{\beta} \neq \emptyset$. Thus,

$$\|(S_{\bar{\beta}}^0)^{-1}A_{\bar{\beta}}^T d_{\bar{\beta}}^0\| \leq \sqrt{\bar{\lambda}_{\bar{\beta}}} \|d_{\bar{\beta}}^0\|. \quad (3.12)$$

Therefore, we have,

$$\begin{aligned} \|(S^0)^{-1}A^T d_\beta^0\|^2 &= \|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\|^2 + \|(S_{\bar{\beta}}^0)^{-1}A_{\bar{\beta}}^T d_{\bar{\beta}}^0\|^2 \\ &= \|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\|^2 \left(1 + \frac{\|(S_{\bar{\beta}}^0)^{-1}A_{\bar{\beta}}^T d_{\bar{\beta}}^0\|^2}{\|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\|^2} \right) \\ &\leq \|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\|^2 \left(1 + \frac{\bar{\lambda}_{\bar{\beta}}}{\lambda_\beta} \right). \end{aligned}$$

Let

$$\gamma_\beta = 1/\sqrt{1 + \frac{\bar{\lambda}_{\bar{\beta}}}{\lambda_\beta}}. \quad (3.13)$$

Then, we have

$$\frac{\|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\|}{\|(S^0)^{-1}A^T d_\beta^0\|} \geq \gamma_\beta. \quad (3.14)$$

If $\bar{\beta} = \emptyset$, the ratio is equal to 1. \square

If we choose γ in (3.6) as,

$$\gamma < \gamma_\beta = 1/\sqrt{1 + \frac{\bar{\lambda}_{\bar{\beta}}}{\lambda_\beta}}, \quad (3.15)$$

then $y(\gamma)$ satisfies

$$\begin{aligned} \|(S^0)^{-1}A^T(y(\gamma) - y^0)\| &= \gamma\|(S^0)^{-1}A^T d_\beta^0\| \\ &\leq \gamma\|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\|/\gamma_\beta = \frac{\gamma}{\gamma_\beta} < 1, \end{aligned} \quad (3.16)$$

since $\|(S_\beta^0)^{-1}A_\beta^T d_\beta^0\| = 1$.
Thus, we have

$$\|(S^0)^{-1}(s(\gamma) - s^0)\| = \|(S^0)^{-1}A^T(y(\gamma) - y^0)\| < 1. \quad (3.17)$$

Hence

$$s(\gamma) > 0.$$

The conclusion is:

Lemma 3 *If $\gamma < \gamma_\beta$, then $(y(\gamma), s(\gamma))$ of (3.6) and (3.9) is an interior feasible dual solution.*

Here, we will try to expand the active set β and show that γ_β of (3.13) is non decreasing with respect to β . For this purpose, we will refer to a lemma (Wilkinson [9]) concerning with the eigenvalues of a symmetric matrix perturbed by a rank-one matrix.

Lemma 4 *Suppose $B = A + \tau cc^T$ where $A \in R^{n \times n}$ is symmetric, $c \in R^n$ has unit 2-norm and $\tau(> 0) \in R$, then*

$$\lambda_i(B) \in [\lambda_i(A), \lambda_{i-1}(A)] \quad (i = 2, \dots, n) \quad (3.18)$$

where $\lambda_i(A)$ ($\lambda_i(B)$) is the i -th largest eigenvalue of A (B).

Moreover, there exist nonnegative m_1, \dots, m_n such that

$$\lambda_i(B) = \lambda_i(A) + m_i\tau \quad (i = 1, \dots, n) \quad (3.19)$$

with $m_1 + \dots + m_n = 1$.

Let the sets β and $\bar{\beta}$ be defined by (3.2) and (3.3). We will try to expand β by adding an element $h \in \bar{\beta}$ and define

$$\beta + 1 = \beta \cup \{h\} \quad (3.20)$$

and

$$\overline{\beta + 1} = \bar{\beta} - \{h\}. \quad (3.21)$$

For $\beta + 1$ and $\overline{\beta + 1}$, we define $A_{\beta+1}$, $A_{\overline{\beta+1}}$, $S_{\beta+1}^0$, $S_{\overline{\beta+1}}^0$, $\Delta_{\beta+1}$, $\bar{\lambda}_{\overline{\beta+1}}$ and $\gamma_{\beta+1}$ in the same way as for β and $\bar{\beta}$.

Lemma 5

$$\gamma_{\beta+1} \geq \gamma_{\beta}. \quad (3.22)$$

Proof. Since

$$A_{\beta+1}(S_{\beta+1}^0)^{-2}A_{\beta+1}^T = A_{\beta}(S_{\beta}^0)^{-2}A_{\beta}^T + s_h^{-2}A_hA_h^T,$$

we have, by Lemma 4,

$$\Delta_{\beta+1} \geq \Delta_{\beta}.$$

On the other hand, since

$$A_{\overline{\beta+1}}(S_{\overline{\beta+1}}^0)^{-2}A_{\overline{\beta+1}}^T = A_{\overline{\beta}}(S_{\overline{\beta}}^0)^{-2}A_{\overline{\beta}}^T + s_h^{-2}A_hA_h^T,$$

we have

$$\bar{\lambda}_{\overline{\beta}} \geq \bar{\lambda}_{\overline{\beta+1}}.$$

Hence,

$$\gamma_{\beta+1} = 1/\sqrt{1 + \frac{\bar{\lambda}_{\overline{\beta+1}}}{\Delta_{\beta+1}}} \geq 1/\sqrt{1 + \frac{\bar{\lambda}_{\overline{\beta}}}{\Delta_{\beta}}} = \gamma_{\beta}. \quad \square$$

If we continue to expand the active set β until $J = \{1, \dots, n\}$, we have the lemma:

Lemma 6 *The lower bound γ_{β} to the ratio $\|(S_{\beta}^0)^{-1}A_{\beta}^T d_{\beta}^0\|/\|(S^0)^{-1}A^T d_{\beta}^0\|$ is monotone nondecreasing with respect to β and the last $\gamma_J (= 1)$ coincides with the ratio.*

4 Primal Variable

Let

$$p_\beta(\bar{z}^0) = (S_\beta^0)^{-1} A_\beta^T \Gamma_\beta^{-1} \nabla \phi_D(y^0, \bar{z}^0), \quad (4.1)$$

with

$$\Gamma_\beta = A_\beta (S_\beta^0)^{-2} A_\beta^T.$$

Then, we have

$$\nabla \phi_D^T(y^0, \bar{z}^0)(y(\gamma) - y^0) = -\gamma \|p_\beta(\bar{z}^0)\|. \quad (4.2)$$

$p_\beta(\bar{z}^0)$ can be rewritten as

$$\begin{aligned} p_\beta(\bar{z}^0) &= -\frac{\rho}{\bar{z}^0 - b^T y^0} (S_\beta^0)^{-1} A_\beta^T \Gamma_\beta^{-1} b + (S_\beta^0)^{-1} A_\beta^T \Gamma_\beta^{-1} A (S^0)^{-1} e \\ &= -\frac{\rho}{\bar{z}^0 - b^T y^0} (S_\beta^0)^{-1} \left[(S_\beta^0)^{-2} A_\beta^T \Gamma_\beta^{-1} b - \frac{\bar{z}^0 - b^T y^0}{\rho} (S_\beta^0)^{-2} A_\beta^T \Gamma_\beta^{-1} A (S^0)^{-1} e \right. \\ &\quad \left. + \frac{\bar{z}^0 - b^T y^0}{\rho} (S_\beta^0)^{-1} e_\beta \right] + e_\beta. \end{aligned} \quad (4.3)$$

Let us define

$$\begin{aligned} x_\beta^1 &= (S_\beta^0)^{-2} A_\beta^T \Gamma_\beta^{-1} b - \frac{\bar{z}^0 - b^T y^0}{\rho} (S_\beta^0)^{-2} A_\beta^T \Gamma_\beta^{-1} A (S^0)^{-1} e \\ &\quad + \frac{\bar{z}^0 - b^T y^0}{\rho} (S_\beta^0)^{-1} e_\beta \\ &= (S_\beta^0)^{-2} A_\beta^T \Gamma_\beta^{-1} \left[b - \frac{\bar{z}^0 - b^T y^0}{\rho} A (S^0)^{-1} e \right] + \frac{\bar{z}^0 - b^T y^0}{\rho} (S_\beta^0)^{-1} e_\beta. \end{aligned} \quad (4.4)$$

Then, we have

$$p_\beta(\bar{z}^0) = -\frac{\rho}{\bar{z}^0 - b^T y^0} (S_\beta^0) x_\beta^1 + e_\beta. \quad (4.5)$$

From (4.4), we have

$$\begin{aligned} A_\beta x_\beta^1 &= b - \frac{\bar{z}^0 - b^T y^0}{\rho} A(S^0)^{-1} e + \frac{\bar{z}^0 - b^T y^0}{\rho} A_\beta (S_\beta^0)^{-1} e_\beta \\ &= b - \frac{\bar{z}^0 - b^T y^0}{\rho} A_{\bar{\beta}} (S_{\bar{\beta}}^0)^{-1} e_{\bar{\beta}}. \end{aligned} \quad (4.6)$$

Let

$$x_{\bar{\beta}}^1 = \frac{\bar{z}^0 - b^T y^0}{\rho} (S_{\bar{\beta}}^0)^{-1} e_{\bar{\beta}}, \quad (4.7)$$

and

$$x^1 = \begin{bmatrix} x_\beta^1 \\ x_{\bar{\beta}}^1 \end{bmatrix}. \quad (4.8)$$

Then, we have

$$Ax^1 = A_\beta x_\beta^1 + A_{\bar{\beta}} x_{\bar{\beta}}^1 = b. \quad (4.9)$$

Thus, x^1 is a solution of $Ax = b$. Furthermore, from (4.7),

$$x_{\bar{\beta}}^1 > 0. \quad (4.10)$$

Let

$$p = \begin{bmatrix} p_\beta(\bar{z}^0) \\ p_{\bar{\beta}}(\bar{z}^0) \end{bmatrix}, \quad (4.11)$$

with

$$p_{\bar{\beta}}(\bar{z}^0) = -\frac{\rho}{\bar{z}^0 - b^T y^0} (S_{\bar{\beta}}^0) x_{\bar{\beta}}^1 + e_{\bar{\beta}} = 0. \quad (\text{by (4.7)}) \quad (4.12)$$

The last relation shows that the $\bar{\beta}$ part of p is always zero and *centered*.

Lemma 7 *Let $\rho = n + \nu\sqrt{n}$ (with $\nu = O(1) > 1$) and $0 < \alpha < 1$. If $\|p\| \leq \alpha$, then x^1 of (4.8) is an interior primal feasible solution and we have*

$$\begin{aligned} (x^1)^T s^0 &= c^T x^1 - b^T y^0 \\ &\leq \frac{n + \alpha\sqrt{n}}{n + \nu\sqrt{n}} (c^T x^0 - b^T y^0) \\ &= \frac{n + \alpha\sqrt{n}}{n + \nu\sqrt{n}} (x^0)^T s^0. \end{aligned}$$

Thus, the duality gap is reduced at least by $(n + \alpha\sqrt{n})/(n + \nu\sqrt{n}) (< 1)$.

Proof. Since

$$\begin{aligned} \|p\| &= \|p_{\bar{\beta}}(\bar{z}^0)\| \\ &= \left\| -\frac{\rho}{\bar{z}^0 - b^T y^0} (S_{\bar{\beta}}^0) x_{\bar{\beta}}^1 + e_{\bar{\beta}} \right\| \leq \alpha < 1, \end{aligned}$$

we have $x_{\bar{\beta}}^1 > 0$. Thus, x^1 is an interior primal feasible solution. On the other hand,

$$\left| e^T \left(-\frac{\rho}{\bar{z}^0 - b^T y^0} (S^0) x^1 + e \right) \right| \leq \|e\| \|p\| \leq \alpha\sqrt{n}$$

and

$$e^T \left(-\frac{\rho}{\bar{z}^0 - b^T y^0} (S^0) x^1 + e \right) = -\frac{\rho}{(x^0)^T s^0} (x^1)^T s^0 + n.$$

Hence,

$$\frac{\rho}{(x^0)^T s^0} (x^1)^T s^0 \leq n + \alpha\sqrt{n}.$$

Noting, $\rho = n + \nu\sqrt{n}$, we have

$$(x^1)^T s^0 \leq \frac{n + \alpha\sqrt{n}}{n + \nu\sqrt{n}} (x^0)^T s^0. \quad \square$$

Ye [10] essentially proved the following lemma which will be used in Section 6:

Lemma 8 (Ye) *If $\|p\| < \alpha < 1/\sqrt{2}$, then we have*

$$n \ln(x^1)^T s^0 - \sum_{j=1}^n \ln(x_j^1 s_j^0) \leq n \ln(x^0)^T s^0 - \sum_{j=1}^n \ln(x_j^0 s_j^0) + \frac{\alpha^2(1 + \alpha/\sqrt{1 - \alpha^2})}{2(1 - 2\alpha^2)}. \quad (4.13)$$

5 Algorithm

One iteration of the dual potential reduction algorithm with the active set strategy is as follows:

Algorithm A

Let $\rho = n + \nu\sqrt{n}$ ($\nu = O(1) > 1$), $\alpha = 0.4$ and $\delta = 0.05$.

Given $s_0 = c - A^T y^0 > 0$, $Ax^0 = b$, $x^0 > 0$ and $\bar{z}^0 = c^T x^0$,

begin

 Define the index set β and $\bar{\beta}$,

 begin inner iteration:

 Compute

x_β^1 of (4.4), and

$p_\beta(\bar{z}^0)$ of (4.5).

 If $\|p_\beta(\bar{z}^0)\| \geq \alpha$ then

 begin dual step:

 Define

$$y(\gamma) = y^0 - \gamma \Gamma_{\beta}^{-1} \left[-\frac{\rho}{\bar{z}^0 - b^T y^0} b + A(S^0)^{-1} e \right] \quad (5.1)$$

and

$$s(\gamma) = c - A^T y(\gamma). \quad (5.2)$$

Then find

$$\gamma^* = \operatorname{argmin}_{\gamma > 0} \phi_D(y(\gamma), \bar{z}^0). \quad (5.3)$$

Let

$$\delta^* = \phi_D(y^0, \bar{z}^0) - \phi_D(y(\gamma^*), \bar{z}^0). \quad (5.4)$$

If $\delta^* < \delta$, then expand β by adding an element of $\bar{\beta}$ with the smallest s_j^0 and redefine β and $\bar{\beta}$.

Go back to the beginning of *inner iteration*.

If $\delta^* \geq \delta$, then end inner iteration and let

$$y^1 = y(\gamma^*), \quad s^1 = s(\gamma^*) \quad (5.5)$$

$$x^1 = x^0, \quad \bar{z}^1 = \bar{z}^0. \quad (5.6)$$

end dual step:

else if $\|p_{\beta}(\bar{z}^0)\| < \alpha$, then begin primal step:

Compute

$$x_{\beta}^1 \text{ of (4.7).}$$

Let

$$x^1 = \begin{bmatrix} x_{\beta}^1 \\ x_{\bar{\beta}}^1 \end{bmatrix} \quad (5.7)$$

$$\bar{z}^1 = c^T x^1 \quad (5.8)$$

$$y^1 = y^0, \quad s^1 = s^0. \quad (5.9)$$

end primal step:

end one iteration.

The iteration will be repeated until $(x^0)^T s^0 < 2^{-L}$ is satisfied.

6 Convergence

Lemma 9 *The inner iteration in Algorithm A ends after at most $(n - m)$ additions to β and comes to satisfy $\delta^* \geq \delta$.*

Proof. In the dual step, we have

$$\|p\| = \|p_\beta(\bar{z}^0)\| \geq \alpha.$$

If we set, in (5.1),

$$\gamma < \gamma_\beta = 1/\sqrt{1 + \frac{\bar{\lambda}_\beta}{\Delta_\beta}}, \quad (6.1)$$

then, by Lemma 3, $(y(\gamma), s(\gamma))$ satisfies (3.16) and (3.17) and is dual feasible and hence by Lemma 1, we have

$$\begin{aligned} \phi_D(y(\gamma), \bar{z}^0) - \phi_D(y^0, \bar{z}^0) &\leq \nabla \phi_D^T(y^0, \bar{z}^0)(y(\gamma) - y^0) + \frac{\|(S^0)^{-1}(s(\gamma) - s^0)\|^2}{2(1 - \|(S^0)^{-1}(s(\gamma) - s^0)\|_\infty)} \\ &\leq -\gamma \|p_\beta(\bar{z}^0)\| + \frac{(\gamma/\gamma_\beta)^2}{2(1 - \gamma/\gamma_\beta)} \\ &\leq -\gamma\alpha + \frac{(\gamma/\gamma_\beta)^2}{2(1 - \gamma/\gamma_\beta)}. \end{aligned} \quad (6.2)$$

If we choose $\alpha = 0.4$ and

$$\gamma = \frac{0.4\gamma_\beta^2}{1 + 0.8\gamma_\beta} (< \gamma_\beta), \quad (6.3)$$

then

$$-\gamma\alpha + \frac{(\gamma/\gamma_\beta)^2}{2(1 - \gamma/\gamma_\beta)} = -\frac{0.08\gamma_\beta^2}{1 + 0.4\gamma_\beta}. \quad (6.4)$$

Let

$$\delta(\beta) = \frac{0.08\gamma_\beta^2}{1 + 0.4\gamma_\beta}. \quad (6.5)$$

Then, $\delta(\beta)$ is monotone nondecreasing with respect to γ_β . Since, by Lemma 5, γ_β is monotone nondecreasing with respect to β and $\gamma_J = 1$, $\delta(\beta)$ is monotone nondecreasing in β and we have

$$\delta(J) = 0.08/1.4 = 0.0572. \quad (6.6)$$

Hence, for $\delta = 0.05$, there is a minimal β which satisfies

$$\delta^* \geq \delta$$

and

$$\phi_D(y(\gamma^*), \bar{z}^0) = \phi_D(y^0, \bar{z}^0) - \delta^* \leq \phi_D(y^0, \bar{z}^0) - \delta. \quad (6.7)$$

□

Lemma 10 *In the preceding Algorithm, either, in the dual step, we have*

$$\phi_{PD}(x^0, s^1) \leq \phi_{PD}(x^0, s^0) - \delta \quad (6.8)$$

or in the primal step, we have

$$\phi_{PD}(x^1, s^0) \leq \phi_{PD}(x^0, s^0) - \delta \quad (6.9)$$

with $\delta = 0.05$.

Proof. In the *dual step*, we have, from (6.7) and (2.11)

$$\phi_{PD}(x^0, s^1) \leq \phi_{PD}(x^0, s^0) - \delta.$$

In the *primal step*, we have by Lemma 8,

$$\begin{aligned}
n \ln(x^1)^T s^0 - \sum_{j=1}^n \ln(x_j^1 s_j^0) &\leq n \ln(x^0)^T s^0 - \sum_{j=1}^n \ln(x_j^0 s_j^0) \\
&+ \frac{\alpha^2(1 + \alpha/\sqrt{1 - \alpha^2})}{2(1 - 2\alpha^2)}. \tag{6.10}
\end{aligned}$$

On the other hand, by Lemma 7, it holds

$$\begin{aligned}
\sqrt{n} (\ln(x^1)^T s^0 - \ln(x^0)^T s^0) &\leq \sqrt{n} \ln \frac{n + \alpha\sqrt{n}}{n + \nu\sqrt{n}} \\
&\leq \ln \frac{1 + \alpha}{1 + \nu} \\
&= \ln \left(1 + \frac{\alpha - \nu}{1 + \nu} \right) \\
&< \frac{\alpha - \nu}{1 + \nu} = -1 + \frac{1 + \alpha}{1 + \nu}. \tag{6.11}
\end{aligned}$$

From (6.10) and (6.11), we have

$$\begin{aligned}
\phi_{PD}(x^1, s^0) &\leq \phi_{PD}(x^0, s^0) - \nu + \frac{\nu(1 + \alpha)}{1 + \nu} \\
&+ \frac{\alpha^2(1 + \alpha/\sqrt{1 - \alpha^2})}{2(1 - 2\alpha^2)} \\
&\leq \phi_{PD}(x^0, s^0) - \frac{1}{2} + \frac{\alpha}{2} \\
&+ \frac{\alpha^2(1 + \alpha/\sqrt{1 - \alpha^2})}{2(1 - 2\alpha^2)},
\end{aligned}$$

(the last inequality holds since $-\nu + \nu(1 + \alpha)/(1 + \nu)$ is monotone decreasing for $\nu \geq 1$.)

If $\alpha = 0.4$, then it holds

$$\phi_{PD}(x^1, s^0) \leq \phi_{PD}(x^0, s^0) - 0.13 \leq \phi_{PD}(x^0, s^0) - \delta. \quad \square$$

The following lemma is proved in Ye [10].

Lemma 11 (Ye) *Let $\rho = n + \nu\sqrt{n}$ with $\nu \geq 1$ and (x^k, s^k) ($k = 0, 1, 2, \dots$) be a series of interior primal-dual feasible pairs with $\phi_{PD}(x^0, s^0) = O(\sqrt{n}L)$. If, for a positive δ independent of n , the relation*

$$\phi_{PD}(x^{k+1}, s^{k+1}) \leq \phi_{PD}(x^k, s^k) - \delta$$

holds for each k , then in $O(\nu\sqrt{n}L)$ iterations, we have

$$c^T x^k - b^T y^k = (x^k)^T s^k < 2^{-L}.$$

If $O(\nu) = 1$ moreover, then the polynomial bound of iteration is $O(\sqrt{n}L)$.

We will roughly estimate the number of arithmetic operations within one iteration. Related with $(A_\beta(S_\beta^0)^{-2}A_\beta^T)^{-1}$, we need $O(n^3)$. In the inner iteration, we can use the rank-one update for modifying $(A_\beta(S_\beta^0)^{-2}A_\beta^T)^{-1}$ after addition of one active constraint, whose cost is $O(n^2)$. In the worst case, we need it $(n-m)$ times and hence its cost is $O(n^3)$. Totally, each iteration uses $O(n^3)$ arithmetic operations.

Now, we have the theorem:

Theorem 1 *If Algorithm A starts from an interior primal-dual feasible solution (x^0, s^0) with $\phi_{PD}(x^0, s^0) = O(\sqrt{n}L)$, then it terminates in $O(\sqrt{n}L)$ iterations and each iteration uses $O(n^3)$ arithmetic operations.*

Remark 6.1

In the inner iteration, we adopt the smallest dual slack policy to expand the active set. If we add the constraint h to β , then by Lemma 4 there exist σ_1

and σ_2 such that

$$\Delta_{\beta+1} = \Delta_{\beta} + s_h^{-2} \sigma_1 \quad (0 \leq \sigma_1 \leq 1)$$

and

$$\bar{\lambda}_{\beta+1} = \bar{\lambda}_{\beta} - s_h^{-2} \sigma_2. \quad (0 \leq \sigma_2 \leq 1)$$

So, it is likely that a certain amount of increase in $\gamma_{\beta+1}$ may be expected by this policy.

Remark 6.2

If we do not assume $\|A_j\| = 1$ ($j = 1, \dots, n$), the ordering (3.1) should be in the ascending order of $s_j^0 / \|A_j\|$. Also, in the inner iteration we use this order in expanding the active set.

Remark 6.3

The choice of $\alpha = 0.4$ and $\delta = 0.05$ is only for theoretical purpose. It is a subject of future study to determine practically efficient α and δ . Also, there may be other active set strategies.

7 A Special Active Set Strategy

Here, we observe a special case when the active set consists of m elements:

$$|\beta| = m. \quad (7.1)$$

Then, A_{β} constitutes a *basis* of A . We denote A_{β} as B and $A_{\bar{\beta}}$ as N . Also, we use the notations x_B or x_N instead of x_{β} or $x_{\bar{\beta}}$.

Using the basis B , we have

$$\Gamma_B^{-1} = (B^T)^{-1} (S^0)^2 B^{-1} \quad (7.2)$$

and we can simplify the formulae for x_B^1 , $p_B(\bar{z}^0)$ and $y(\gamma)$ as follows:

$$x_B^1 = B^{-1} \bar{b} + \tau (S_B^0)^{-1} e_B \quad (7.3)$$

$$p_B(\bar{z}^0) = -\frac{1}{\tau} (S_B^0) x_B^1 + e_B \quad (7.4)$$

and

$$y(\gamma) = y^0 - \gamma(B^T)^{-1}S_B^0 p_B(\bar{z}^0), \quad (7.5)$$

where

$$\bar{b} = b - \tau A(S^0)^{-1}e \quad (7.6)$$

and

$$\tau = \frac{\bar{z}^0 - b^T y^0}{\rho} \quad (7.7)$$

are fixed within one iteration.

We can solve linear equations related with B^{-1} and $(B^T)^{-1}$ by using efficient methods such as the *LU* decomposition of B .

In the inner iteration, let the added active constraints be A_{h_1}, \dots, A_{h_k} and

$$\Gamma_{B+k} = B(S_B^0)^{-2}B^T + \sum_{i=1}^k s_{h_i}^{-2}A_{h_i}A_{h_i}^T. \quad (7.8)$$

Then, the formulae for the expanded active set are as follows:

$$x_{B+k}^1 = (S_{B+k}^0)^{-2}B_k\Gamma_{B+k}^{-1}\bar{b} + \tau(S_{B+k}^0)^{-1}e_{B+k} \quad (7.9)$$

$$p_{B+k}(\bar{z}^0) = -\frac{1}{\tau}(S_{B+k}^0)x_{B+k}^1 + e_{B+k} \quad (7.10)$$

and

$$y(\gamma) = y^0 + \frac{\gamma}{\tau}\Gamma_{B+k}^{-1}\bar{b}, \quad (7.11)$$

where

$$B_k = [B, A_{h_1}, \dots, A_{h_k}] \in R^{m \times (m+k)}. \quad (7.12)$$

The key factor for this step is the computation of $\Gamma_{B+k}^{-1}\bar{b}$. For this purpose, let $U_i \in R^m$ be the solution of

$$B(S_B^0)^{-1}U_i = s_{h_i}^{-1}A_{h_i}. \quad (i = 1, \dots, k) \quad (7.13)$$

What we need is to deal with B^{-1} , $(B^T)^{-1}$ and a sequence of rank-one modifications of type (7.17). This may bring us a big economy in the numerical implementation of the algorithm. We can use several efficient techniques developed for the simplex method ([2], [1]).

In order to make $|\beta| = m$, it is sufficient to change the active set strategy in Section 3 as follows:

1. Let us reorder the index set $J = \{1, \dots, n\}$ in the ascending order of s_j^0 :

$$s_1^0 \leq s_2^0 \leq \dots \leq s_n^0.$$

2. Take linearly independent m vectors from A_1, A_2, \dots, A_n in this order. Thus, we have the index sets $\beta = \{j_1, \dots, j_m\}$ and $\bar{\beta} = J - \beta$.

As to the basis change, we can use the same basis for several iterations so long as the potential reduction is big. However, if we use the same basis for a long time, the ellipsoid constraint

$$\|(S^0)^{-1}A^T(y - y^0)\| \leq \gamma \tag{7.21}$$

may come to be very flat in the direction $y(\gamma)$. As a result, a short step, a small potential reduction and hence many additions to the active set will come out. There may be a trade-off between the frequency of basis change and the amount of works.

8 Conclusion

In this paper, we proposed an active set strategy for Ye's dual potential reduction algorithm. This strategy allows us to deal with $(A_\beta(S_\beta^0)^{-2}A_\beta^T)^{-1}$ instead of $(A(S^0)^{-2}A^T)^{-1}$, thus resulting in reduction of computational time when there are many redundant constraints in the dual optimal solution. Then, we observed the special case: $|\beta| = m$, where we can be free from the burden of $(A_\beta(S_\beta^0)^{-2}A_\beta^T)^{-1}$ and deal only with B^{-1} , $(B^T)^{-1}$ and rank-one modifications of inverse matrices. In this case, several efficient techniques of the simplex method are available. It is a subject of future research to see the

effects of this strategy by numerical experiments.

Acknowledgement. I would like to thank Franco Giannessi for the talks which motivated this work.

References

- [1] V.Chvátal, *Linear Programming*, Freeman, 1983.
- [2] G.B.Dantzig, *Linear Programming and Extensions*, Princeton University Press, 1963.
- [3] R.M.Freund, "Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function," Operations Research Center, Massachusetts Institute of Technology, Cambridge, 1988.
- [4] C.Gonzaga, "An algorithm for solving linear programming problems in $O(n^3L)$ operations," in: N.Megiddo, ed., *Advances in Mathematical Programming-Interior Point and Related Methods*, Springer Verlag, 1989.
- [5] C.Gonzaga, "Large-steps path following methods for linear programming: potential reduction method," Department of Systems Engineering and Computer Science, COPPE-Federal University of Rio of Janeiro, Brasil, 1989.
- [6] N.Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica* 4(1984)373-395.
- [7] M.Kojima, S.Mizuno and A.Yoshise, "A primal-dual interior point algorithm for linear programming," in: N.Megiddo, ed., *Advances in Mathematical Programming-Interior Point and Related Methods*, Springer Verlag, 1989.
- [8] M.J.Todd and Y.Ye, "A centered projective algorithm for linear programming," Technical Report 763, School of ORIE, Cornell University, 1987.
- [9] J.H.Wilkinson, *The Algebraic Eigenvalue Problem*, Claredon Press, 1965.
- [10] Y.Ye, "An $O(n^3L)$ potential reduction algorithm for linear programming," Dept. of Management Sciences, The University of Iowa, 1989.