Algorithmic Operations Research Vol.7 (2013) 51–54

Fast Descent Methods for LPs With No Matrix Inversions

Katta G. Murty [a]

[a]Department of IOE, University of Michigan, Ann Arbor, MI 48109-2117, USA;

Abstract

Existing software implementations for solving Linear Programming (LP) models are all based on full matrix inversion operations involving every constraint in the model in every step. This **linear algebra component** *in these systems makes it difficult to solve dense models even with moderate size, and it is also the source of accumulating roundoff errors affecting the accuracy of the output.*

We present a new Sphere method, SM-6, for LP not using any pivot steps. The method is currently undergoing computational tests.

Key words: Linear Programming (LP), Interior point methods (IPMs) , solving LPs by descent methods without using matrix inversions.

In Memorium: I dedicate this paper to the memory of my dear friend Santosh Kabadi with whom I had many fruitful discussions on the methods discussed in this paper, who passed away in a tragic drowning accident in the sacred Ganges river recently.

1. Introduction

For modeling decision making applications, LP is the most commonly used mathematical model. Software systems for solving LP models are based on either the simplex method, or interior point methods (IPMs, in particular the primal-dual IPM) developed during the second half of the 20th century (see the books [1–3,6] for detailed discussion of these methods, and references on them) and are able to solve large scale sparse models (those involving thousands of constraints), within reasonable times by exploiting the sparcity of the models. As several real world applications lead to sparse models, these systems are very popular in practice.

But the simplex method, and these IPMs are based on matrix inversion operations involving every constraint in the model in every step. In large scale applications, these matrix inversion operations limit the ability of these algorithms to only those with very sparse coefficient matrices. Typically, the effectiveness of these algorithms fades as the density of the coefficient matrix increases.

Many applications lead to LP models that are not sparse, and need near-optimum solutions in real time. In many of these applications, the LP models are only of moderate size. This provided a motivation for us to develop fast algorithms for LP without using matrix inversion operations.

SMs consider LPs in the form:

$$\begin{array}{rl} \min & z = cx \\ \text{s.t.} & Ax \geq b \end{array} \qquad (1)$$

where A is an $m \times n$ data matrix; with a known interior feasible solution x^1 (i.e., satisfying $Ax^1 > b$). Here is the notation we will use in this paper.

- **Notation for rows and columns of A:** $A_{i.}, A_{.j}$ denote the i^{th} row and j^{th} column of A.
- **Feasible region and its interior:** K denotes the set of feasible solutions of (1), and $K^0 = \{x : Ax > b\}$ its interior.
- **Facetal hyperplanes:** $FH_i = \{x : A_{i.}x = b_i\}$, the i-th facetal hyperplance of K for i = 1 to m.
- **Largest inscribed ball with a given point as center, its radius:** $B(x, \delta(x))$, $\delta(x)$ are defined for $x \in K^0$. $\delta(x)$ = minimum$\{\frac{A_{i.}x-b_i}{||A_{i.}||} : i = 1, ..., m\}$ is the radius of the largest ball that can be inscribed in K with x as its center. $B(x, \delta(x)) = \{y : ||y - x|| \leq \delta(x)\}$ is that largest inscribed ball in K with x as its center. We will use " $B(x)$" to denote the ball

Email: Katta G. Murty [murty@umich.edu].

$B(x, \delta(x))$.

- **Touching constraint set at a given point:** $T(x)$ defined for $x \in K^0$, it is the set of all indices i satisfying: $\frac{A_{i.}x - b_i}{||A_{i.}||}$ = Minimum$\{\frac{A_{p.}x - b_p}{||A_{p.}||} : p = 1$ to $m\} = \delta(x)$. The facetal hyperplane $FH_i = \{x : A_{i.}x = b_i\}$ is a tangent plane to $B(x, \delta(x))$ for each $i \in T(x)$, that's why $T(x)$ is called the **index set of touching constraints in (1) defining** K, at x.
- **GPTC (gradient projection on touching constraint) directions:** Let c^i denote the orthogonal projection of c^T on $\{x : A_{i.}x = 0\}$, i.e., $c^i = c^T - A_{i.}^T[(A_{i.}c^T/||A_{i.}||^2]$ for i = 1 to m. When the ball $B(x, \delta(x))$ is under consideration, the directions $-c^i$ for $i \in T(x)$ are called the **GPTC directions at the current center** x in K.
- **Ball center of** K**:** It is the center of a largest ball in K, it maximizes $\delta(x)$ over $x \in K$.
- **IFS:** We will use this abbreviation for "interior feasible solution", i.e., for K it will be an x satisfying $Ax > b$.
- **For** $\hat{x} \in K$**,** $H(\hat{x})$**:** is the objective plane through $\hat{x}$, it is $\{x : cx = c\hat{x}\}$.

Sphere Methods (SM) for LP were introduced by Murty in 2006 and developed further (see Chapter 8 on Sphere Methods in the book [4], and paper [5]) . In this paper we will describe SM-6 for LP, which does not use any pivot steps at all.

2. Sphere Method 6 (SM-6) for LP

In this method, in a general **descent step** from an IFS x^* in descent direction d (i.e., d satisfying $cd < 0$), we move from x^* in this direction the maximum distance possible while still remaining at a distance ϵ from the boundary. This gives the **step length** in this descent step to be β, where

$$\beta = \text{Minimum } \{(-A_{i.}x^* + b_i + \epsilon)/(A_{i.}d) : \text{over all } i \text{ satisfying } A_{i.}d < 0\} \qquad (2)$$

if there is at least one i satisfying $A_{i.}d < 0$, and in this case the output of this descent step is $x^* + \beta d$. Here ϵ is a small positive tolerance, like 0.1. In this case the best value for ϵ wiil be determined from computational tests.

On the other hand, if $A_{i.}d \geq 0$ for all i, then the step length in this descent step is ∞. In this case the feasible half-line $\{x^* + \beta d : \beta \geq 0\}$ is one along which the objective function cx in (1) diverges to $-\infty$ as β goes to ∞. If this occurs in any descent step carried out in this method, we terminate the algorithm with the conclusion that cx is **unbounded below** in (1)

Subroutine 1: Finding the interval of values of a real parameter α say, satisfying a given system of linear inequalities: In every iteration in SM-6, we encounter the problem of finding the interval of values of a single variable α satisfying a given system of linear inequalities on it. Here we give a procedure for this problem. Let the system be :

$$a_t + d_t\alpha \geq 0 \text{ for all } t = 1 \text{ to } k$$

We assume that $d_t \neq 0$ for all t, as otherwise that inequality does not involve α at all. The required interval is: $\bar{\alpha}^1$ = max$\{(-a_t/d_t)$: for all t satisfying $d_t > 0\} \leq \alpha \leq \bar{\alpha}^2 =$ min$\{(-a_t/d_t$: for all t satisfying $d_t < 0\}$.

Here define maximum [minimum] in the empty set to be $-\infty$ $[+\infty]$ respectively. If $\bar{\alpha}_1 > \bar{\alpha}_2$ then the system has no solution.

Each iteration in SM-6 begins with an initial IFS, and terminates by producing an output IFS at the end.. The first iteration begins with the IFS x^1 given in input data. Subsequent iterations begin with the output IFS of the previous iteration. Each iteration consists of several steps. We will describe each of these steps now. In the rest of this Section 2 we will describe the algorithm under the assumption that K is **bounded**. We will now describe the steps in a general iteration in SM-6 under this assumption. The case whre K may not be bounded will be discussed in the next Section 3.

The General Iteration in SM-6

Let $\bar{x}$ denote the initial IFS with which this iteration begins.

Step C1: Finding an approximate center beginning with $\bar{x}$:

Substep 1. Move the objective plane $H(\bar{x})$ in the direction $-c^T$ until it becomes a tangent plane to the ball $B(\bar{x})$, the largest ball inside K with $\bar{x}$ as center (with its radius $\delta(\bar{x})$). The point where the objective plane in its new position touches $B(\bar{x})$ is $\bar{x}^1 = \bar{x} - c^T[\delta(\bar{x})/||c||]$.

If $\bar{x}^1$ is a boundary point of K, i.e., satisfies $A_{i.}\bar{x}^1 = b_i$ for at least one $i \in \{1, ..., m\}$, then this $\bar{x}^1$ is an optimum solution of the LP (1), terminate the algorithm. Otherwise continue. Now apply the following Substep 2 for each $i \in T(\bar{x})$:

Substep 2. For $i \in T(\bar{x})$, x^i = point where the factal hyperplane FH_i touches $B(\bar{x})$. So it is the orthogonal projection of $\bar{x}$ on FH_i, therefore $x^i = \bar{x} - A_{i.}^T(\delta(\bar{x})/||A_{i.}||) = \bar{x} - A_{i.}^T[(A_{i.}\bar{x} - b_i)/||A_{i.}||^2]$.

For each $i \in T(\bar{x})$ compute the orthogonal projection x^{i1} of x^i on $H(\bar{x}^1)$, $x^{i1} = x^i - c^T[(cx^i - c\bar{x}^1)/cc^T]$. Define L_i = straight line joining x^{i1} and $\bar{x}^1$, it is $\{x^{i1}(\gamma) = x^{i1} + \gamma(\bar{x}^1 - x^{i1}) : \gamma$ takes all real values$\}$, in parametric form in terms of parameter γ. To determine the interval of values of the parameter γ for which $x^{i1}(\gamma)$ lies in K, we need to solve the following system of linear inequalities in γ.

$$A_{p.}x^{i1}(\gamma) \geq b_p \text{ for } p = 1 \text{ to } m \qquad (3)$$

This is a system of m linear inequalities in the parameter γ, and we know that when $\gamma = 1$, $x^{i1}(\gamma) = \bar{x}^1$ which is an interior point of K. So the interval of values of (3) is a nonempty interval with $\gamma = 1$ as an interior point.

(3) can be solved by Subroutine 1 discussed above, and since we assumed that K is bounded in this discussion, the interval of solutions will be a bounded interval $\gamma_{i1} \leq \gamma \leq \gamma_{i2}$. Find γ_{i1}, γ_{i2} using Subroutine 1. Then:

$L_i \cap K = \{ x^{i1}(\gamma) : \gamma_{i1} \leq \gamma \leq \gamma_{i2}\}$, and the length of this interval , denoted by ℓ_i is $||x^{i1}(\gamma_{i1}) - x^{i1}(\gamma_{i2})||$.

Substep 3. Among all $i \in T(\bar{x})$ select that index i which corresponds to the maximum value for ℓ_i computed in Substep 2, and suppose it is $i = r$. Then the output point $\bar{\bar{x}}^1$ from this Step C1 beginning with $\bar{x}$, is $\bar{\bar{x}}^1 = [x^{r1}(\gamma_{r1}) + x^{r1}(\gamma_{r2})]/2$, the mid-point of $L_r \cap K$.

Actually, in this Substep 3, there is an alternate procedure for selecting the output point $\bar{\bar{x}}^1$. For this procedure let $\Delta = \{(x^{i1}(\gamma_{i1}) + x^{i1}(\gamma_{i2}))/2 : i \in T(\bar{x})\}$, the set of mid-points of the line segments $L_i \cap K$ for $i \in T(\bar{x})$. Under this alternative, take $\bar{\bar{x}}^1$ as the point $x \in \Delta$ corresponding to the maximum value for $\delta(x)$. We will determine which of these alternative procedures in Substep 3 gives better results, and select that one as the procedure to implement in Substep 3.

Step C2: Finding an approximate center beginning with $\bar{\bar{x}}^1$, the output point from Step C1

Carry out Step C1 beginning with $\bar{\bar{x}}^1$ instead of $\bar{x}$, and suppose the output point obtained is $\bar{\bar{x}}^2$. This point $\bar{\bar{x}}^2$ is called the **Center** in this iteration. Now go to the Descent steps in this iteration.

Descent steps in this iteration

Select a set Γ, initially = $\emptyset$, for storing the output points with their objective values (for $z = cx$) generated in each of the descent steps carried out below.

Descent step D5.1: The current center is $\bar{\bar{x}}^2$. For each $i \in T(\bar{\bar{x}}^2)$, compute the **touching point** x^i of $B(\bar{\bar{x}}^2)$ with FH_i, $x^i = \bar{\bar{x}}^2 - A_{i.}^T[(A_{i.}\bar{\bar{x}}^2 - b_i)/||A_{i.}||^2]$. Then $\hat{x}^i = \epsilon\bar{\bar{x}}^2 + (1-\epsilon)x^i$ is called the **NTP (Near touching point)** of $B(\bar{\bar{x}}^2)$ with FH_i. It is the point ϵ distance away from x^i on the line segment joining x^i to $\bar{\bar{x}}^2$, where ϵ is a small positive tolerance.

For each $i \in T(\bar{\bar{x}}^2)$, take a descent step from the NTP $\hat{x}^i$ in the descent direction $-c^i$. Store the output points from each of these descent steps along with their objective values in the set Γ.

Other descent steps: From the center $\bar{\bar{x}}^2$, take descent steps in the directions $\bar{\bar{x}}^2 - \bar{\bar{x}}^1$, $-c^T$, and the average of the GPTC directions at the current center $\bar{\bar{x}}^2$. Store the output points from each of these descent steps along with the objective values at them, in the set Γ.

After all these descent steps, select the point in Γ corresponding to the least objective value as the new $\bar{x}$, reset the set Γ to be the empty set, and with the new $\bar{x}$ go to the next iteration.

Terminate the method when the change in objective value in an iteration falls below a selected tolerance. In the final iteration, take the point in Γ corresponding to the least objective value as an approximate optimum solution of (1).

3. What to do if we don't know whether K is bounded

Suppose we do not know whether K is bounded or not. In this case also, we apply the algorithm as described above. Now several cases may occur.

Case 1: The algorithm may continue as usual, with all γ_{i1}, γ_{i2} finite until termination with an approximate optimum solution at the end. In this case we get that approximate optimum for (1) without knowing whether K is bounded or not.

Case 2: The algorithm may continue as usual, with all γ_{i1}, γ_{i2} finite in every step, until in some descent step, the step legth turns out to be ∞, and we terminate with the conclusion that cx is unbounded below in (1). That final descent line provides a feasible half-line along which cx diverges to $-\infty$.

Case 3: In some iteration, in Substep 2, either $\gamma_{i1} = -\infty$, or $\gamma_{i2} = \infty$ for some i in the touching set $T(\bar{x})$ in that iteration.

In this Substep consider the line $L_i = \{x^{i1}(\gamma) = x^{i1} + \gamma(\bar{x}^1 - x^{i1}) : \gamma$ takes real values$\}$ for that i. We consider the two possibilities separately.

Possibility 1: Suppose $\gamma_{i1} = -\infty$. Then $A_{p.}(\bar{x}^1 - x^{i1}) \leq 0$ for all p = 1 to m.

If $A_{p.}(\bar{x}^1 - x^{i1}) < 0$ for all $p = 1$ to m, $\delta(x^{i1} + \gamma(\bar{x}^1 - x^{i1}))$ diverges to ∞ as γ goes to $-\infty$, so under this possibility, cx diverges to $-\infty$, i.e., it is unbounded below over K.

If $A_{p.}(\bar{x}^1 - x^{i1}) \leq 0$ for all p = 1 to m, and is equal to 0 for some p, then move to the descent steps to take descent steps from $x^{i1}(\gamma)$ for some $\gamma < 0$, and continue the algorithm by going to the next iteration with the output point at the end of these descent steps.

Possibility 2: Suppose $\gamma_{i1} = \infty$. Then $A_{p.}(\bar{x}^1 - x^{i1}) \geq 0$ for all p = 1 to m.

If $A_{p.}(\bar{x}^1 - x^{i1}) > 0$ for all $p = 1$ to m, $\delta(x^{i1} + \gamma(\bar{x}^1 - x^{i1}))$ diverges to ∞ as γ goes to ∞, so under this possibility, cx diverges to $-\infty$, i.e., it is unbounded below over K.

If $A_{p.}(\bar{x}^1 - x^{i1}) \geq 0$ for all p = 1 to m, and is equal to 0 for some p, then move to the descent steps to take descent steps from $x^{i1}(\gamma)$ for some $\gamma > 0$, and continue the algorithm by going to the next iteration with the output point at the end of these descent steps.

References

[1] G. B. Dantzig and M. N. Thapa, *Linear Programming, Vol. 1. Introduction , Vol. 2. Theory and Extensions*: Springer-Verlag New York. 1997.

[2] N. Megiddo, Pathways to the optimal set in linear programming, *Progress in Mathematical Programming: Interior Point and Related Methods*, N. Meggiddo, ed., Springer-Verlag, New York, ch. 8, (1989) 131-158.

[3] R. D. C. Monteiro and I. Adler, Interior path-following primal-dual algorithms, Part I: Linear programming, Mathematical Programming, 44(1989) 27-41.

[4] K. G. Murty, *Optimization for Decision Making: Linear and Quadratic Models*, Springer, NY, 2009.

[5] K. G. Murty, and M. R. Oskoorouchi, Sphere Methods for LP, *Algorithmic Operations Research*, 5(2010) 21-33.

[6] S. J. Wright, *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, PA, 1997.

Received 2-1-2013; accepted 14-2-2013