Ye's Interior Point Algorithm.

CS270 Algorithms & Data Structures **Spring 2008**

Lecture 13 Ye's Interior Point Algorithm. : 3.04.08

Lecturer: Satish Rao *Scribe: Rao, for now*

Disclaimer: *These are rough notes, with some exercises.*

Question: Have we shown how to solve linear programs in polynomial time?

No. The simplex algorithm was not proven to work in this time.

Question: Do we know if there is a polynomial time algorithm?

Sure. The ellipsoid algorithm does so. We won't get into that today. But it is especially useful in theory. Since you are not (all) in theory, we will look at the approach that competes with simplex in practice and is still polynomial time: Interior point methods.

Question: Must be optimizing in the interior. Can we eventually get a vertex solution?

Sure. If one is very close to a vertex, one can get to a nearby vertex.

Question: First, how close can two vertices be?

Well, you can examine the solution of linear systems (again) since vertices are such things. Cramer's rule gives one an explicit solution. And the denominators have a number of bits. This gives a bound on the precision required, and in fact, on the minimum distance between two distinct vertices. The bound is on the order of 2^{-L} where L is the total number of bits in the input. (Yuck, I know, but so be it, for now)

Question: Can we get from a point inside that is very close to a vertex to a vertex?

Yes, but we won't discuss it further today.

Question: Which form of linear program will we work with today?

$$minz = cx$$

$$AX = b$$
$$x \geq 0$$

and its dual

$$maxw = by$$
$$A^T y + s = c$$
$$s \geq 0$$

We will simultaneously solve both. That is, it is a primal-dual algorithm. We will do so by maintaining a pair of solutions (x, s) that are interior to the polytope.

Question: Why stay in the interior?

To remain feasible. (It skirts the question of how to start, but so be it, for now.)

Question: What does interior mean anyway?

$$x > 0, s > 0.$$

Question: What is a pair optimal?

$$x^T s = 0.$$

Question: How can I transform the coordinate so that at least the x's are away from the border of their polytope?

The following scaling of for a vector $\bar{x}$ should work: x_i is mapped to $x_i/\bar{x}_i$. This maps the vector $\bar{x}$ to $e = (1, ..., 1)$. That is, all the x_i are clearly away from 0.

Now, we can rewrite the primal dual problem with respect to some scaling $\bar{x}$ replacing A with $\bar{A} = \bar{X}A$ where $\bar{X}$ is the diagonal matrix with $\bar{x}$ as its diagonal.

Now, the solution to the resulting dual s' would need to be scaled with respect to $\bar{x}$ as follows; $s_j = \bar{x}_j s'_j$ to get a solution to the original problem.

Question: What is a pair optimal?

The optimality condition remains unchanged since the scalings cancel. That is,

$$x^T s = x'^T s'.$$

Question: So some algorithms do the obvious. What is the obvious?

Greedy. So, we should try to greedily move toward optimal but stay in interior.

Question: How do we combine two goals?

Let's use a potential function. The following one...

$$G(x, s) = q \ln(x^T s) - \sum_{j=1}^{n} \ln(x_j s_j)$$

The first term goes negative as we approach optimality. The second goes positive as the x_j or s_j go toward 0.

In particular, $q = n + \sqrt{n}$. (We remark that for $q = n$ it is always positive, i.e., $n \log n$, so q better be bigger than n.)

Exercise 1: Prove that for $q = n$, that the potential function is always greater than $n \ln n$? Exercise 2: When all the terms $x_j s_j$ are equal, what is the potential function for $q = n + l$? Exercise 3: When one of the terms $x_j s_j$ is less than the average by a factor

of f (and the others are equal) , what is the potential function for $q = n + l$?

Question: When can we stop?

When $x^T s \leq 2^{-2L}$, since at this point, we can jump to an optimal vertex.

Question: We wish to drive down the potential function. How low?

Lemma 13.1 *If x, s are feasible primal-dual vectors $G(x, s) \leq -k\sqrt{n}L$, then $x^T s < e^{-kL}$.*

Proof:

$$-k\sqrt{n}L \geq G(x,s) = (n + \sqrt{n}) \ln x^T s - \sum_j \ln x_j s_j.$$

This is

$$\geq \sqrt{n} \ln x^T s + n \ln n.$$

And

$$\ln x^T s \leq -kL - \sqrt{n} \ln n.$$

■

Question: How high is $G(x, s)$ at the start?

Well, one needs to show it is $O(\sqrt{n}L)$. This occurs as long as we begin with a feasible solution with all the x_i, s_i's are greater than $1/2^L$. This can be accomplished by using a different linear program and it is messy; we leave it for another day.

Question: Now what?

$i = 0$, Start with $x^0, s^0 > 0$ and y^0 such that $Ax^0 = b, A^T y^0 + s^0 = c$ and $G(x^0, y^0) = O(\sqrt{n}L$.

Question: Continue?

while $G(x^i, s^i)$ is large, do a step to get x^{i+1}, s^{i+1}.

Question: Whats a step?

Improve $G(e, s')$ by movining along its gradient in the transformed space! Move some direction. Let's move either x or s.

Question: What direction?

Reduce $G()$! Take some derivatives.

$$g = \Delta_x G(x, s)|_{(e,s')}.$$

This is,

$$\frac{q}{x^T s} s - 1/x_i|_{(e,s')}.$$

or in the direction

$$g = \frac{q}{e^T s'} s' - e.$$

Question: Stay feasible?

Project onto null space of $\overline{A}$ to find a vector d.

Claim: $d = (I - \bar{A}(\bar{A}\bar{A}^T)^{-1})g$.

This follows from the facts that $\bar{A}d = 0$, and $\bar{A}^T w = g - d$. Then solve for w and get

$$(\bar{A}\bar{A}^T)w = \bar{A}(g - d) = \bar{A}g$$

$$d = g - (g - d) = g - A^T w.$$

Plugging in, we get the claim.

Question: What now?

If d is large, then good. That is, if $||d|| \geq .4$

Update as follows...

$$\tilde{x} = e - \frac{1}{4||d||} d.$$

Question: Are the $\tilde{x}_i' s$ positive?

Sure, we are subtracting something less than 1 from each which used to be ok.

Question: Do we improve?

Will show later that we improve the potential function by $-7/120$.

Question: What if $||d||$ is small?

Then, we make a dual step. Perhaps take a derivative of G with respect to s.

$$h = \frac{q}{e^T s'} e - 1/\bar{s}_i'.$$

Notice $h_i = g_i/s_i$. Thus we will punt and work with g since it is at least related and hope for the best.

Question: We want to make a feasible move in direction g in the dual. How?

Let's look at the equations..

$$\bar{A}^T y' + s' = \bar{c}$$

For any y

$$\bar{A}^T y + \tilde{s} = \bar{c}$$

where $\tilde{s} - s' = \bar{A}^T(y' - y)$.

Thus, we move perpendicular to the null space and in the direction $-(g - d)$.

Thus, we have

$$\tilde{(s)} = s' - (g - d)\mu$$

for some μ.

We choose $\mu = \frac{e^T s'}{q}$ and can get $\bar{A}^T(y' + \mu w) + \tilde{s} = c$.

Thus, we get that

$$\tilde{s} = s' - \frac{e^T s'}{q}(g - d).$$

Plugging in for $(g - d)$, we get

$$\tilde{s} = \frac{e^T s'}{q}(d + e)$$

Question: Are $\tilde{s}$'s positive?

Yes. You will show this.

Question: Does this help?

We will improve the potential by $-1/6$.

Question: How many iterations?

$O(\sqrt{n}L)$.

Question: How long?

Well, we need to compute d. This requires an inversion of $(\bar{A}\bar{A}^T)$, which can be done in $O(n^3)$. An approximation is sufficient and this can be done with time $O(n^{2.5})$.