A New Iteration-Complexity Bound for the MTY Predictor-Corrector Algorithm

Renato Monteiro (Georgia Tech)
Takashi Tsuchiya (Inst. of Stat. Math.)

McMaster Univ, Hamilton CA

March 10, 2003

A New Iteration-Complexity Bound for the MTY Predictor-Corrector Algorithm

Renato Monteiro (Georgia Tech)
Takashi Tsuchiya (Inst. of Stat. Math.)

IMA Workshop on Semidefinite Programming and Robust Optimization

March 14, 2003

Talk Outline

- related works;
- LP problem and assumptions;
- central path and its neighborhood;
- Mizuno-Todd-Ye predictor-corrector (MTY P-C) algorithm;
- condition number and scale-invariance;
- new iteration-complexity bounds for finding;
  - a near optimal solution;
  - an exact optimal solution (with the use of a scaling-invariant FT procedure);
- brief outline of our analysis;
- concluding remarks.

Related Works

The Vavasis-Ye layered-step method is an interior-point method for solving the LP problem $\min\{c^T x : Ax = b,\ x \geq 0\}$, where $A \in \Re^{m\times n}$. Using the notion of crossover events, they showed that their algorithm has an iteration-complexity which depends neither on b nor c, namely

$$\mathcal{O}\left(n^{3.5}\log(\bar{\chi}_A + n)\right),$$

where $\bar{\chi}_A$ is a condition number for A. ($\bar{\chi}_A$ is known to be 2^{L_A}, where L_A is the input size of A.) In addition to the ordinary steps of path following IPMs, their algorithm uses layered least squares (LLS) steps from time to time.

One drawback of their algorithm is that an estimate of $\bar{\chi}_A$, which is very difficult to obtain, is needed for computing the LLS step.

Megiddo, Mizuno and Tsuchiya (Math Progr 98) and Monteiro and Tsuchiya (SIOPT) have proposed distinct ways to get around this problem.

Related Works (continued)

The Mizuno-Todd-Ye (MTY) predictor-corrector (P-C) algorithm is a popular and theoretically appealing method which converges both polynomially and quadratically.

Using the notion of crossover events too, we derive a new iteration-complexity bound for the MTY P-C algorithm in terms of the condition number $\bar{\chi}_A$.

The new iteration complexity bound is scaling invariant, and becomes

$$\mathcal{O}(n^{3.5} L_A + n^2 \log L))$$

under the Turing Machine Model, where L is the input size of (A, b, c) and n is the number of variables. This contrasts strongly with the best classical bound of $\mathcal{O}(\sqrt{n}\, L)$, which is proportional to L instead of $\log L$.

The LP problem

$$
\begin{array}{lll}
(P) & \text{minimize}_x & c^T x \\
& \text{subject to} & Ax = b,\ x \geq 0,
\end{array}
$$

$$
\begin{array}{lll}
(D) & \text{maximize}_{(y,s)} & b^T y \\
& \text{subject to} & A^T y + s = c,\ s \geq 0,
\end{array}
$$

Assumptions

1) (P) and (D) have interior-feasible solutions.
2) the rows of the $m \times n$ matrix A are linearly independent.

Definition The duality gap of a feasible (x, y, s) is given by

$$c^T x - b^T y = (A^T y + s)^T x - b^T y = x^T s.$$

Central path and its neighborhood

For each $\nu > 0$, the system

$$\begin{aligned} XSe &= \nu e, \\ Ax - b &= 0, \qquad (x,s) \geq 0, \\ A^T y + s - c &= 0, \end{aligned}$$

where $X = \text{Diag}(x)$, $S = \text{Diag}(s)$ and $e = (1, \ldots, 1)^T$, has a unique solution $(x(\nu), y(\nu), s(\nu))$, which converges to a primal-dual optimal solution as $\nu \to 0$.

The MTY P-C is based on the 2-norm neighborhood of the central path:

$$\mathcal{N}(\beta) \equiv \{(x, y, s) \text{ feasible} : \|Xs - \mu e\| \leq \beta\mu\},$$

where $\mu = \mu(x, s) \equiv (x^T s)/n$ and β is a positive constant.

MTY P-C algorithm

Search directions: For a feasible (x, y, s), the Newton direction towards the point $(x(\nu), y(\nu), s(\nu))$ is obtained by solving the linear system:

$$\begin{aligned} X\Delta s + S\Delta x &= -Xs + \nu e, \\ A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0 \end{aligned}$$

Setting $\nu = 0$ yields the predictor (or affine scaling) direction at (x, y, s). This direction aims towards an optimal solution of (P) and (D).

Setting $\nu = \mu(x, s)$ yields the corrector (or centrality) direction at (x, y, s). This direction aims towards the point on the central-path having the same duality gap as the current point (x, y, s).

An iteration of the MTY P-C Alg.

Let $w = (x, y, s) \in \mathcal{N}(\beta)$ be given, where $\beta \in (0, 1/4]$.

1) Compute the AS direction $\Delta w^a = (\Delta x^a, \Delta y^a, \Delta s^a)$ at w;

2) Let $\alpha_p > 0$ be the maximum $\alpha \in [0, 1]$ such that $w + \alpha \Delta w^a \in \mathcal{N}(2\beta)$;

3) Set $w_p = w + \alpha_p \Delta w^a$;

4) Compute the corrector direction $\Delta w^c = (\Delta x^c, \Delta y^c, \Delta s^c)$ at w_p;

5) The next point w^+ is determined as $w^+ = w_p + \Delta w^c$;

It can be proved that $w^+ \in \mathcal{N}(\beta)$. Hence, a new iteration can be started by setting $w \leftarrow w^+$ and going back to 1).

Classical convergence results

Lemma: Suppose $w = (x, y, s) \in \mathcal{N}(\beta)$, where $\beta \in (0, 1/4]$. Then, $w^+ \in \mathcal{N}(\beta)$ and

$$\frac{\mu(w^+)}{\mu(w)} \leq \min\left\{1 - \sqrt{\frac{\beta}{n}}, \frac{\varepsilon_\infty^{\mathrm{a}} \sqrt{n}}{\beta}\right\},$$

where $\delta = s^{1/2} x^{-1/2}$ and

$$\varepsilon_\infty^{\mathrm{a}} \equiv \max_i \left\{\min\left\{\left|\frac{\delta_i^{-1} \Delta s_i^a}{\sqrt{\mu}}\right|, \left|\frac{\delta_i \Delta x_i^a}{\sqrt{\mu}}\right|\right\}\right\}.$$

Consequences:

- The first bound implies that the MTY algorithm converges polynomially, namely in $\mathcal{O}(\sqrt{n}\,L)$ iterations under the TMM.
- Since $\varepsilon_\infty^{\mathrm{a}} = \mathcal{O}(\mu(w))$ asymptotically, the second bound implies that $\mu(w^+) = \mathcal{O}(\mu(w)^2)$. This implies that the MTY algorithm converges quadratically.

The condition number $\bar{\chi}_A$

Define

$$\bar{\chi}_A \equiv \sup\{\|(ADA^T)^{-1}AD\| : D \in \mathcal{D}\},$$

where $\mathcal{D}$ denotes the set of all positive definite diagonal matrices.

Facts:

1) if A integral then $\bar{\chi}_A \leq 2^{L_A}$, where L_A is the input size of A.

2) Finding an upper bound for $\bar{\chi}_A$ is a $\mathcal{NP}$ hard problem.

3) $\bar{\chi}_A = \max\{\|A_B^{-1}A\| : A_B \text{ is a basis of } A\}$.

Scale Invariance

Let D be a positive diagonal matrix and consider the pair of LPs:

$$
\begin{array}{lll}
(\tilde{P}) & \text{minimize}_{\tilde{x}} & (Dc)^T \tilde{x} \\
 & \text{subject to} & AD\tilde{x} = b, \;\; \tilde{x} \geq 0, \\
 & & \\
(\tilde{D}) & \text{maximize}_{(\tilde{y},\tilde{s})} & b^T \tilde{y} \\
 & \text{subject to} & DA^T \tilde{y} + \tilde{s} = \tilde{c}. \;\; \tilde{s} \geq 0,
\end{array}
$$

obtained from (P) and (D) by performing the change of variables $(x, y, s) = \Phi(\tilde{x}, \tilde{y}, \tilde{s}) \equiv (D\tilde{x}, \tilde{y}, D^{-1}\tilde{s})$.

The MTY P-C algorithm is scaling-invariant, i.e., if $\{w^k\}$ and $\{\tilde{w}^k\}$ denote the sequence of iterates generated by the MTY P-C algorithm in the original and the scaled space, then $w^k = \Phi(\tilde{w}^k)$ for all $k \geq 1$, as long as the initial iterates satisfy this relation too, that is $w^0 = \Phi(\tilde{w}^0)$.

Remark: Vavasis-Ye algorithm and its variants are not scaling invariant.

New Complexity Results

Theorem 1: Given $\eta > 0$ and an initial point $w^0 \in \mathcal{N}(\beta)$, with $\beta \in (0, 1/4]$, the MTY P-C algorithm generates an iterate $w^k \in \mathcal{N}(\beta)$ satisfying $\mu(w^k) \leq \eta$ in at most

$$\mathcal{O}\left(\min\left(\begin{array}{c}\sqrt{n}\,\log(\mu_0/\eta)\\ n^2\log(\log(\mu_0/\eta)) + n^{3.5}\log(\bar{\chi}_A^* + n)\end{array}\right)\right)$$

iterations, where $\mu_0 \equiv \mu(w^0)$ and

$$\bar{\chi}_A^* \equiv \inf\{\bar{\chi}_{AD} : D \in \mathcal{D}\}.$$

Remark: $\bar{\chi}_A^*$ is a scaling invariant condition number of A.

Finite termination procedure

Given a feasible point $w = (x, y, s)$, it consists of two steps:

- determine a guess $(B(w), N(w))$ for the optimal partition (B_*, N_*) defined as

$$
\begin{aligned}
B_* &\equiv \{i : x_i^* > 0 \text{ for some } x^* \in \text{opt}(P)\},\\
N_* &\equiv \{i : s_i^* > 0 \text{ for some } (y^*, s^*) \in \text{opt}(D)\}.
\end{aligned}
$$

- "project" w onto the "manifold" determined by $(B(w), N(w))$.

The partition $(B, N) = (B(w), N(w))$ is found as

$$
\begin{aligned}
B &= \left\{ i : \left| \frac{\delta_i \Delta x_i^a}{\sqrt{\mu}} \right| \le \left| \frac{\delta_i^{-1} \Delta s_i^a}{\sqrt{\mu}} \right| \right\},\\
N &= \{1, \ldots, n\} \setminus B,
\end{aligned}
$$

where $\delta \equiv s^{1/2} x^{-1/2}$.

Remark: This partition is scaling-invariant.

Next, we solve the following projection problems:

$$x^* \equiv \operatorname{argmin}_{\tilde{x}} \left\{ \|\delta(x - \tilde{x})\|^2 : A\tilde{x} = b,\ \tilde{x}_N = 0 \right\},$$

$$(y^*, s^*) \equiv \operatorname{argmin}_{(\tilde{y}, \tilde{s})} \left\{ \|\delta^{-1}(s - \tilde{s})\|^2 : \begin{array}{c} A^T\tilde{y} + \tilde{s} = c \\ \tilde{s}_B = 0 \end{array} \right\}.$$

If $x^*_B > 0$ and $s^*_N > 0$ then $w^* = (x^*, y^*, s^*)$ is a strictly complementary optimal solution; output w^* and declare success. Otherwise, exit the procedure and declare failure.

Proposition: There exists a scaling invariant constant $\eta_* > 0$ such that the above FT Procedure always succeeds whenever $\mu(w) \leq \eta_*$.

Theorem 2: Suppose $w^0 \in \mathcal{N}(\beta)$, where $\beta \in (0, 1/4]$. Then, the MTY P-C algorithm with the above FT Procedure started from w^0 finds a primal-dual strictly complementary optimal solution w^* in at most

$$\mathcal{O}\left(\min\left\{\begin{array}{c}\sqrt{n}\,\log(n\mu_0/\eta_*)\\ n^2\log(\log(n\mu_0/\eta_*)) + n^{3.5}\log(\bar{\chi}_A^* + n)\end{array}\right\}\right)$$

iterations, where $\mu_0 = \mu(w^0)$, and

$$\eta_* \equiv \sup\left\{\frac{[\xi(AD, b, Dc)]^2}{\zeta(AD,(B_*,N_*))} : D \in \mathcal{D}\right\}.$$

By definition, $\xi(A, b, c) \equiv \min\left\{\min_{i\in B_*}\xi_i^P, \min_{i\in N_*}\xi_i^D\right\}$, where for every $i = 1, \ldots, n$,

$$\begin{aligned}\xi_i^P &\equiv \max\{\bar{x}_i : \bar{x} \text{ optimal for (P)}\},\\ \xi_i^D &\equiv \max\{\bar{s}_i : (\bar{y}, \bar{s}) \text{ optimal for (D)}\}.\end{aligned}$$

Also, $\zeta(A, (B_*, N_*)) \equiv \max\{\zeta_P, \zeta_D\}$, where

$$\zeta_P \equiv \max_{d_{N_*}\neq 0}\left\{\min\left\{\frac{\|d_{B_*}\|}{\|d_{N_*}\|} : d = (d_{B_*}, d_{N_*}) \in \operatorname{Ker}(A)\right\}\right\},$$

$$\zeta_D \equiv \max_{d_{B_*}\neq 0}\left\{\min\left\{\frac{\|d_{N_*}\|}{\|d_{B_*}\|} : d = (d_{B_*}, d_{N_*}) \in \operatorname{Im}(A^T)\right\}\right\}.$$

Complexity bounds under the TMM

Assume that the data (A, b, c) is integral. Let L and L_A denote the input size of (A, b, c) and A, respectively.

Using the big M idea, Vavasis and Ye constructed an auxiliary pair of LP problems with the following properties:

i) the input size of its coefficient matrix is bounded by $\mathcal{O}(L_A)$;

ii) its cost and right hand coefficients are bounded by $\mathcal{O}(L)$;

iii) it admits a readily available well-centered initial point whose duality gap is $n\,2^{\mathcal{O}(L)}$, and;

iv) if (P) and (D) have an optimal solution, then such a solution can be easily obtained from an optimal solution of the auxiliary pair of LPs.

Theorem 3: Suppose that (P) and (D) have a primal-dual optimal solution. Then, the MTY P-C algorithm together with the FT Procedure applied to the above auxiliary pair of LP problems, finds a strictly complementary primal-dual optimal solution of (P) and (D) in

$$\mathcal{O}\left(\min\left\{\sqrt{n}\,L\,,\, n^2\log L + n^{3.5}(L_A + \log n)\right\}\right)$$

iterations.

Brief Analysis Outline

Definition (Vavasis and Ye): For a constant $\mathcal{C} \geq 1$, a $\mathcal{C}$-crossover event is said to occur on the interval $(\nu_1, \nu_2]$ if there exists indices $i, j \in \{1, \ldots, n\}$ such that

$$s_j(\nu_2) \leq \mathcal{C}\, s_i(\nu_2), \text{ and}$$

$$s_j(\nu) > \mathcal{C}\, s_i(\nu) \text{ for all } \nu \leq \nu_1.$$

Proposition: For any fixed $\mathcal{C} \geq 1$, there exists at most $n(n-1)/2$ disjoint intervals of the form $(\nu_1, \nu_2]$ containing $\mathcal{C}$-crossover events.

.

- Basic result of Vavasis and Ye on the relationship between LLS steps and crossover events.
- Proximity result between the primal-dual affine scaling direction and LLS steps;
- A new partition introduced by Monteiro and Tsuchiya;
- Realization that the MTY P-C algorithm at times (i.e., $\mathcal{O}(n^2)$ times) behaves quadratically convergent.

$$\Downarrow$$

There exists $\mathcal{C} > 0$ such that a $\mathcal{C}$-crossover event occurs every

$$\mathcal{O}\left(\log(\log(\mu_0/\eta)) + n^{1.5}\log(\bar{\chi}_A + n)\right)$$

iterations. Since there are at most $\mathcal{O}(n^2)$ disjoint $\mathcal{C}$-crossover events, the MTY P-C algorithm terminates in

$$\mathcal{O}\left(n^2\log(\log(\mu_0/\eta)) + n^{3.5}\log(\bar{\chi}_A + n)\right).$$

Concluding Remarks

We have presented a new iteration-complexity bound for the MTY P-C algorithm obtained by means of the notion of crossover events.

Possible open problems are:

- Can a general pair of LP problems be solved by the MTY P-C algorithm with the same complexity derived in this paper under the real-computation model?
- Is there an alternative notion that can be used in place of crossover event which, instead of the latter, is scaling-invariant?
- Can an iteration-complexity bound depending only on n and $\bar{\chi}^*_A$ be derived for the MTY P-C algorithm?