Interior Point Methods in Large Neighborhoods of the Central Path:

Florian A. Potra

potra@math.umbc.edu

Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, MD 21250, USA

MOPTA05

Overview

Interior-point methods in mathematical programming have been the largest and most dramatic area of research in optimization since the development of the simplex method. . . Interior-point methods have permanently changed the landscape of mathematical programming theory, practice and computation. . . (Freund & Mizuno 1996).

Major impacts on

- The linear programming problem (LP)
- The quadratic programming problem (QP)
- The linear complementarity problem (LCP)
- The semi-definite programming problem (SDP)
- Some classes of convex programming problems
- Nonlinear Programming

The linear programming problem

$$\min_x c^T x \text{ s.t. } Ax = b, \ x \geq 0.$$

Dantzig (1947-1951): the simplex method

– good practical performance

– exponential worst case complexity (Klee and Minty (1972))

The linear programming problem

$$\min_x c^T x \text{ s.t. } Ax = b, \ x \geq 0.$$

Dantzig (1947-1951): the simplex method

– good practical performance

– exponential worst case complexity (Klee and Minty (1972))

Question: Is (LP) solvable in polynomial time?

(in terms of: L = bitlength of data, and $n = dim(x)$)

The linear programming problem

$$\min_x c^T x \text{ s.t. } Ax = b,\ x \geq 0.$$

Dantzig (1947-1951): the simplex method

– good practical performance

– exponential worst case complexity (Klee and Minty (1972))

Question: Is (LP) solvable in polynomial time?

(in terms of: L = bitlength of data, and $n = dim(x)$)

Answer: YES! Khachyan 1979.

The linear programming problem

$$\min_x c^T x \text{ s.t. } Ax = b,\ x \geq 0.$$

Dantzig (1947-1951): the simplex method
– good practical performance
– exponential worst case complexity (Klee and Minty (1972))

Question: Is (LP) solvable in polynomial time?
(in terms of: L = bitlength of data, and $n = dim(x)$)

Answer: YES! Khachyan 1979.

Proof: The ellipsoid method (an interior point method)

Complexity

Complexity

Khachyan (1979–1980): the ellipsoid method

– iteration complexity: $O(n^2 L)$

– computational complexity: $O(n^4 L)$

Note: Bad practical performance.

Observed complexity same as worst case complexity.

Complexity

Khachyan (1979–1980): the ellipsoid method
– iteration complexity: $O(n^2 L)$
– computational complexity: $O(n^4 L)$
Note: Bad practical performance.
Observed complexity same as worst case complexity.

Karmarkar (1984): projective scaling
– iteration complexity: $O(nL)$
– computational complexity: $O(n^{3.5} L)$
Note: A variant of the algorithm implemented in KORBX (1989).
Practical performance better than theoretical complexity.

Complexity

Khachyan (1979–1980): the ellipsoid method
– iteration complexity: $O(n^2 L)$
– computational complexity: $O(n^4 L)$
Note: Bad practical performance.
Observed complexity same as worst case complexity.

Karmarkar (1984): projective scaling
– iteration complexity: $O(nL)$
– computational complexity: $O(n^{3.5} L)$
Note: A variant of the algorithm implemented in KORBX (1989).
Practical performance better than theoretical complexity.

State of the Art:
– iteration complexity: $O(\sqrt{n}L)$
– computational complexity: $O(n^3 L)$
(Anstreicher (1999) $O(\frac{n^3}{\log n} L)$)
Note: Excellent software packages: CPLEX, LOQO, Mosek, OSL, PCx

How is polynomiality proved?

Primal-dual algorithms:

Obtain a sequence of points with duality gap $\mu_k \to 0$.

Best complexity is obtained for path-following methods where (μ_k) is Q-linearly convergent with Q-factor $(1 - \nu/\sqrt(n))$:

$$\mu_{k+1} \le (1 - \nu/\sqrt{n})\, \mu_k, \quad k = 0, 1, \ldots$$

or for potential reduction method where (μ_k) is R-linearly convergent with R-factor $(1 - \gamma/\sqrt{n})$:

$$\mu_k \le \chi(1 - \nu/\sqrt{n})^k, \quad k = 0, 1, \ldots$$

In both cases we have

$$\mu_k \le \epsilon \;\; for \;\; k = O(\sqrt{n}\log(\frac{\mu_0}{\epsilon}))$$

If $\epsilon \le 2^{-2L}$ then (x^k, y^k, s^k) can be rounded to an exact solution in $O(n^3)$ arithmetic operations. Hence the $O(\sqrt{n}L)$-iteration complexity.

Superlinear convergence

Polynomiality is proved by showing that μ_k converges linearly. Efficient algorithms have superlinear convergence:

$$\mu_{k+1} \leq \alpha_k \, \mu_k, \quad k = 0, 1, \ldots, \qquad \alpha_k \to 0$$

Superlinear convergence

Polynomiality is proved by showing that μ_k converges linearly. Efficient algorithms have superlinear convergence:

$$\mu_{k+1} \leq \alpha_k \, \mu_k, \quad k = 0, 1, \dots, \qquad \alpha_k \to 0$$

Q-quadratic convergence: $\mu_{k+1} \leq \alpha \, \mu_k^2$

Superlinear convergence

Polynomiality is proved by showing that μ_k converges linearly. Efficient algorithms have superlinear convergence:

$$\mu_{k+1} \leq \alpha_k \, \mu_k , \quad k = 0, 1, \ldots , \qquad \alpha_k \to 0$$

Q-quadratic convergence: $\mu_{k+1} \leq \alpha \, \mu_k^2$

superlinear convergence of Q-order ω: $\mu_{k+1} \leq \alpha \, \mu_k^\omega$

Superlinear convergence

Polynomiality is proved by showing that μ_k converges linearly. Efficient algorithms have superlinear convergence:

$$\mu_{k+1} \le \alpha_k \, \mu_k, \quad k = 0, 1, \ldots, \qquad \alpha_k \to 0$$

Q-quadratic convergence: $\mu_{k+1} \le \alpha \, \mu_k^2$

superlinear convergence of Q-order ω: $\mu_{k+1} \le \alpha \, \mu_k^\omega$

- Zhang, Tapia and Dennis (1992): sufficient conditions for superlinear convergence for path-following methods for LP;
- Zhang, Tapia and P. (1993): generalization for QP and LCP;
- Ye, Güler, Tapia and Zhang (1993): The Mizuno-Todd-Ye predictor-corrector method has $O(\sqrt{n}L)$ complexity and Q-quadratic convergence under general conditions;
- Ye and Anstreicher (1993): generalization for LCP under strict complementarity.

Benefits of superlinear convergence

- Convergence is faster than indicated by worst case complexity.
- The condition of the linear systems to be solved at each iteration worsens as μ_k decreases. Superlinearly convergent algorithms need only a couple of iterations with small μ_k.

Benefits of superlinear convergence

- Convergence is faster than indicated by worst case complexity.
- The condition of the linear systems to be solved at each iteration worsens as μ_k decreases. Superlinearly convergent algorithms need only a couple of iterations with small μ_k.

Superlinear convergence vs. finite termination

Mehrotra and Ye (1993): if μ_k is small enough then a projection can be used to find the exact solution. Does superlinear convergence set in (much) before such a projection works?

Benefits of superlinear convergence

- Convergence is faster than indicated by worst case complexity.
- The condition of the linear systems to be solved at each iteration worsens as μ_k decreases. Superlinearly convergent algorithms need only a couple of iterations with small μ_k.

Superlinear convergence vs. finite termination

Mehrotra and Ye (1993): if μ_k is small enough then a projection can be used to find the exact solution. Does superlinear convergence set in (much) before such a projection works?

Superlinear convergence is even more important for SDP since:

- no analogous of simplex
- interior point methods – the only efficient solvers
- no finite termination schemes
- condition of linear systems is more critical

Convergence of the iterates

$\mu_k \to 0$. Is the sequence of iterates convergent?

Convergence of the iterates

$\mu_k \to 0$. Is the sequence of iterates convergent?

- Tapia, Zhang and Ye (1995): sufficient conditions for convergence of iterates. Can they be satisfied with preservation of polynomial complexity?
- Gonzaga and Tapia (1997): the iterates of MTY converge to the analytic center of the solution set.
  MTY: 2 matrix factorizations + 2 backsolves, $\mathcal{Q}(\mu_k) = 2$.
  simplified MTY: asymptotically one factorization + 2 backsolves; the iterates converge but not to the analytic center.
- Bonnans and Gonzaga (1996) Bonnans and P. (1997): General Convergence Theory
- P. (2001): established superlinear convergence of the iterates for several interior point methods.
  The simplified MTY is Q-quadratically convergent.
  MTY appears not to be Q-superlinearly convergent.

Worst case vs. probabilistic analysis

The simplex method has exponential worst case complexity but has good practical performance. Why?

Worst case vs. probabilistic analysis

The simplex method has exponential worst case complexity but has good practical performance. Why?

The expected number of pivots is $O(n^2)$!

(Adler, Borgwardt, Megiddo, Shamir, Smale, Todd; See Shamir(1992))

This is a strongly polynomial probabilistic complexity result
(no dependence on L).

Are there corresponding results for interior point methods?

Worst case vs. probabilistic analysis

The simplex method has exponential worst case complexity but has good practical performance. Why?

The expected number of pivots is $O(n^2)$!

(Adler, Borgwardt, Megiddo, Shamir, Smale, Todd; See Shamir(1992))

This is a strongly polynomial probabilistic complexity result (no dependence on L).

Are there corresponding results for interior point methods?

Anstreicher, Ji, P. and Ye (1999): The expected value of the number of iterations needed for an infeasible start interior point method of MTY type to find an exact solution of the LP or to determine that the LP is infeasible is at most $O(\sqrt{n}\ln n)$!

Worst case vs. probabilistic analysis

The simplex method has exponential worst case complexity but has good practical performance. Why?

The expected number of pivots is $O(n^2)$!

(Adler, Borgwardt, Megiddo, Shamir, Smale, Todd; See Shamir(1992))

This is a strongly polynomial probabilistic complexity result (no dependence on L).

Are there corresponding results for interior point methods?

Anstreicher, Ji, P. and Ye (1999): The expected value of the number of iterations needed for an infeasible start interior point method of MTY type to find an exact solution of the LP or to determine that the LP is infeasible is at most $O(\sqrt{n}\ln n)$!

Numerical experience with a very large number of problems show that interior point methods need 30–50 iterations to convergence. Can the above probabilistic results be improved?

Polylog probabilistic complexity?

Wide neighborhoods of the central path

Primal-dual path-following algorithms acting in wide neighborhoods of the central path are the most efficient interior point methods.
Paradoxically the best complexity results were obtained for algorithms acting in narrow neighborhoods.
Recent work closes this gap (P., Math. Prog. 2003).

Wide neighborhoods of the central path

Primal-dual path-following algorithms acting in wide neighborhoods of the central path are the most efficient interior point methods. Paradoxically the best complexity results were obtained for algorithms acting in narrow neighborhoods.
Recent work closes this gap (P., Math. Prog. 2003).

The horizontal linear complementarity problem (HLCP)

$$\begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\geq 0. \end{aligned}$$

Wide neighborhoods of the central path

Primal-dual path-following algorithms acting in wide neighborhoods of the central path are the most efficient interior point methods.
Paradoxically the best complexity results were obtained for algorithms acting in narrow neighborhoods.
Recent work closes this gap (P., Math. Prog. 2003).

The horizontal linear complementarity problem (HLCP)

$$\begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\geq 0. \end{aligned}$$

HLCP is monotone iff $Qu + Rv = 0 \Rightarrow u^T v \geq 0, \forall u, v \in \mathbb{R}^n$.

Wide neighborhoods of the central path

Primal-dual path-following algorithms acting in wide neighborhoods of the central path are the most efficient interior point methods. Paradoxically the best complexity results were obtained for algorithms acting in narrow neighborhoods.
Recent work closes this gap (P., Math. Prog. 2003).

The horizontal linear complementarity problem (HLCP)

$$\begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\geq 0. \end{aligned}$$

HLCP is monotone iff $Qu + Rv = 0 \Rightarrow u^T v \geq 0, \forall u, v \in \mathbb{R}^n$.

HLCP is skew-symmetric iff $Qu + Rv = 0 \Rightarrow u^T v = 0, \forall u, v \in \mathbb{R}^n$.

Wide neighborhoods of the central path

Primal-dual path-following algorithms acting in wide neighborhoods of the central path are the most efficient interior point methods. Paradoxically the best complexity results were obtained for algorithms acting in narrow neighborhoods.

Recent work closes this gap (P., Math. Prog. 2003).

The horizontal linear complementarity problem (HLCP)

$$\begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\geq 0. \end{aligned}$$

HLCP is monotone iff $Qu + Rv = 0 \Rightarrow u^T v \geq 0, \forall u, v \in I\!R^n$.

HLCP is skew-symmetric iff $Qu + Rv = 0 \Rightarrow u^T v = 0, \forall u, v \in I\!R^n$.

QP reduces to a monotone HLPC. LP reduces to a skew-symmetric HLCP.

The Central Path

The set of all feasible points is denoted by

$$\mathcal{F} = \{z = \lceil x,\, s \rfloor \in I\!R^{2n}_{+} \;\; : \quad Qx + Rs = b\},$$

where $\lceil x,\, s \rfloor$ denotes the vector $[x^T, s^T]^T$. If the relative interior of $\mathcal{F}$,

$$\mathcal{F}^0 = \mathcal{F} \bigcap I\!R^{2n}_{++}$$

is not empty, then the nonlinear system,

$$F_\tau(z) := \begin{bmatrix} xs - \tau e \\ Qx + Rs - b \end{bmatrix} = 0\,.$$

has a unique positive solution for any $\tau > 0$. The set of all such solutions defines the central path $\mathcal{C}$ of the HLCP, i.e.,

$$\mathcal{C} = \{z \in I\!R^{2n}_{++} \;:\; F_\tau(z) = 0,\, \tau > 0\}\,.$$

Central Path Neighborhoods

Proximity measures:

$$\delta_2(z) := \left\| \frac{xs}{\mu(z)} - e \right\|_2, \ \delta_\infty(z) := \left\| \frac{xs}{\mu(z)} - e \right\|_\infty,$$

$$\delta_\infty^-(z) := \left\| \left[\frac{xs}{\mu(z)} - e \right]^- \right\|_\infty, \quad \mu(z) = \frac{x^T s}{n}.$$

Corresponding neighborhoods:

$$\mathcal{N}_2(\alpha) = \{ z \in \mathcal{F}^0 \ : \ \delta_2(z) \le \alpha \},$$

$$\mathcal{N}_\infty(\alpha) = \{ z \in \mathcal{F}^0 \ : \ \delta_\infty(z) \le \alpha \},$$

$$\mathcal{N}_\infty^-(\alpha) = \{ z \in \mathcal{F}^0 \ : \ \delta_\infty^-(z) \le \alpha \}.$$

Relation between neighborhoods:

$$\mathcal{N}_2(\alpha) \subset \mathcal{N}_\infty(\alpha) \subset \mathcal{N}_\infty^-(\alpha), \text{ and } \lim_{\alpha \uparrow 1} \mathcal{N}_\infty^-(\alpha) = \mathcal{F}.$$

First order predictor

Input: $z \in \mathcal{D}(\beta) = \{z \in \mathcal{F}^0 \ : \ xs \geq \beta\mu(z)\} = \mathcal{N}_\infty^-(1-\beta)$.

$$z(\theta) = z + \theta w\,,$$

where

$$w = \lceil\, u,\, v \,\rfloor = -F_0'(z)^{-1} F_0(z)$$

is the Newton direction of F_0 at z (the affine scaling direction), i.e.,

$$\begin{aligned} su + xv &= -xs \\ Qu + Rv &= 0\,. \end{aligned}$$

$$\gamma := \frac{2\,(1-\beta)}{2(1-\,\beta) + n + \sqrt{4(1-\beta)(n+1) + n^2}}\,,$$

$$\overline{\theta} = \text{argmin}\,\{\mu(\theta) \ : \ z(\theta) \in \mathcal{D}((1-\gamma)\beta)\ \}\ ,$$

Output: $\overline{z} = z(\overline{\theta}\,) \in \mathcal{D}((1-\gamma)\beta)$.

First order corrector

Input: $\overline{z} \in \mathcal{D}((1-\gamma)\beta)$.

$$\overline{z}(\theta) = \overline{z} + \theta\overline{w} \ ,$$

where

$$\overline{w} = \lceil \overline{x}, \overline{s} \rfloor = -F'_{\mu(\overline{z})}(\overline{z})^{-1} F_{\mu(\overline{z})}(\overline{z}) \ ,$$

is the centering direction at $\overline{z}$, i.e.,

$$\begin{aligned} \overline{su} + \overline{xv} &= \mu(\overline{z}) - \overline{xs} \\ Q\overline{u} + R\overline{v} &= 0 \, . \end{aligned}$$

Output: $z^+ = \overline{z}(\theta_+) \in \mathcal{D}(\beta)$.

First order predictor–corrector.
Input: $z \in \mathcal{D}(\beta)$, *Output:* $z^+ \in \mathcal{D}(\beta)$.

Iterative algorithm: Given $z^0 \in \mathcal{D}(\beta)$.
For $k = 0, 1, \ldots$: $z \leftarrow z^k; \quad z^{k+1} \leftarrow z^+ , \ \ \mu_{k+1} \leftarrow \mu(z^+) \, , \ k \leftarrow k+1$.

Convergence results

Theorem If HLCP is monotone then the algorithm is well defined and

$$\mu_{k+1} \leq \left(1 - \frac{29\sqrt{(1-\beta)\beta}}{16(n+2)}\right)\mu_k, \quad , k = 0, 1, \ldots$$

Corollary $O(nL)$-iteration complexity.
Theorem If the HLCP has a strictly complementary solution, then the sequence $\{\mu_k\}$ generated by Algorithm 1 converges quadratically to zero in the sense that

$$\mu_{k+1} = O(\mu_k^2)\,.$$

Comments: Same complexity as Gonzaga (1999) plus quadratic convergence. In Gonzaga's algorithm a predictor is followed by an apriori unknown number of correctors. The complexity result is proved by showing that the total number of correctors is at most $O(nL)$. The structure of Gonzaga's algorithm makes it very difficult to analyze the asymptotic convergence properties of the duality gap. No superlinear convergence results have been obtained so far for his method.

A higher order predictor

Input: $z \in \mathcal{D}(\beta)$.

$$z(\theta) = z + \sum_{i=1}^{m} w^i \theta^i \,,$$

where $w^i = \lceil\, u^i,\, v^i \,\rfloor$ are given by

$$\begin{cases} su^1 + xv^1 & = -xs \\ Qu^1 + Rv^1 & = 0 \end{cases} \quad ,$$

$$\begin{cases} su^i + xv^i & = -\sum_{j=1}^{i-1} u^j v^{i-j} \\ Qu^i + Rv^i & = 0 \end{cases} \quad , \quad i = 2, 3, \ldots, m$$

(one factorization + m backsolves ,$O(n^3)$ + $m\, O(n^2)$ arith. operations)

$$\overline{\theta} \; = \text{argmin} \left\{ \mu(\theta) \;:\; z(\theta) \in \mathcal{D}((1-\gamma)\beta) \right\} \; .$$

Output: $\overline{z} = z(\overline{\theta}\,) \in \mathcal{D}((1-\gamma)\beta)\,.$

The higher order predictor corrector

Given $z^0 \in \mathcal{D}(\beta)$.

For $k = 0, 1, \ldots$:

$z \leftarrow z^k$;

Obtain $\bar{z}$ by the higher order predictor;

Obtain z^+ by the first order corrector;

$z^{k+1} \leftarrow z^+ , \;\; \mu_{k+1} \leftarrow \mu(z^+) , \; k \leftarrow k+1.$

Theorem If HLCP is monotone then the algorithm is well defined and

$$\mu_{k+1} \leq \left(1 - .16 \frac{\sqrt{\beta}\,\sqrt[3]{1-\beta)}}{\sqrt{n}\;\sqrt[m+1]{n+2}}\right) \mu_k \,, \quad k = 0, 1, \ldots$$

Corollary $O\left(n^{1/2+1/(m+1)} L\right)$ -iteration complexity.

Corollary If $m = O(\lceil (n+2)^{\omega} - 1 \rceil)$, then $O\left(\sqrt{n} L\right)$ -iteration complexity.

The higher order predictor corrector

Given $z^0 \in \mathcal{D}(\beta)$.
For $k = 0, 1, \ldots$:
$z \leftarrow z^k$;
Obtain $\overline{z}$ by the higher order predictor;
Obtain z^+ by the first order corrector;
$z^{k+1} \leftarrow z^+ \,, \;\; \mu_{k+1} \leftarrow \mu(z^+) \,, \; k \leftarrow k+1.$

Theorem If HLCP is monotone then the algorithm is well defined and

$$\mu_{k+1} \leq \left(1 - .16 \frac{\sqrt{\beta}\,\sqrt[3]{1-\beta)}}{\sqrt{n}\,\sqrt[m+1]{n+2}}\right) \mu_k \,, \quad k = 0, 1, \ldots$$

Corollary $O\left(n^{1/2 + 1/(m+1)} L\right)$ -iteration complexity.

Corollary If $m = O(\lceil (n+2)^{\omega} - 1 \rceil)$, then $O\left(\sqrt{n} L\right)$ -iteration complexity.

$$\lim n^{(1/n^{\omega})} = 1, \quad n^{(1/n^{\omega})} \leq e^{1/(\omega e)}, \; \forall n.$$

The higher order predictor corrector – continued

Theorem We have

$$\mu_{k+1} = O(\mu_k^{m+1}) \text{ , if HLCP is nondegenerate ,}$$

and

$$\mu_{k+1} = O(\mu_k^{(m+1)/2}) \text{ , if HLCP is degenerate .}$$

Conclusion The first algorithm with $O\left(\sqrt{n}L\right)$ -iteration complexity and superlinear convergence for degenerate LCP in the wide neighborhood of the central path.

Remark If we take $\omega = 0.1$ then the values of $m = \lceil (n+2)^{\omega} - 1 \rceil$, corresponding to $n = 10^6$, $n = 10^7$, $n = 10^8$, and $n = 10^9$ are 3, 5, 6, and 7 respectively. This corresponds with efficient practical implementation of interior point methods where the same factorization is used from 3 to 7 times.

The higher order predictor corrector – continued

Theorem We have

$$\mu_{k+1} = O(\mu_k^{m+1}) \text{ , if HLCP is nondegenerate ,}$$

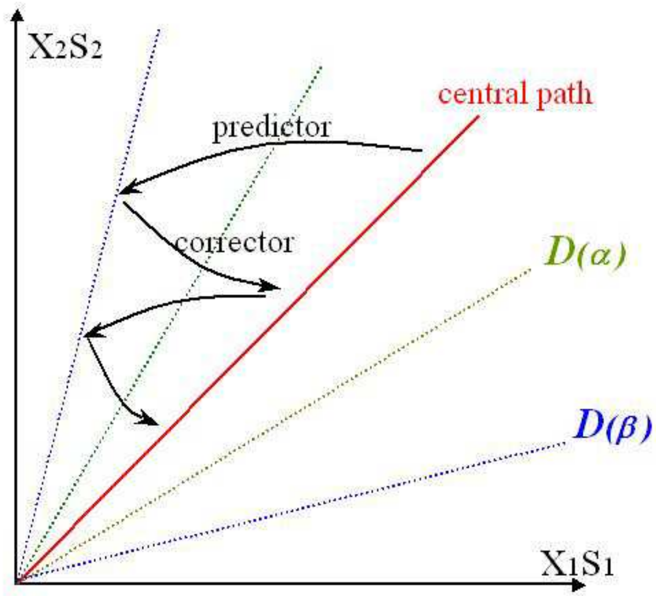
and

$$\mu_{k+1} = O(\mu_k^{(m+1)/2}) \text{ , if HLCP is degenerate .}$$

Conclusion The first algorithm with $O\left(\sqrt{n}L\right)$ -iteration complexity and superlinear convergence for degenerate LCP in the wide neighborhood of the central path.

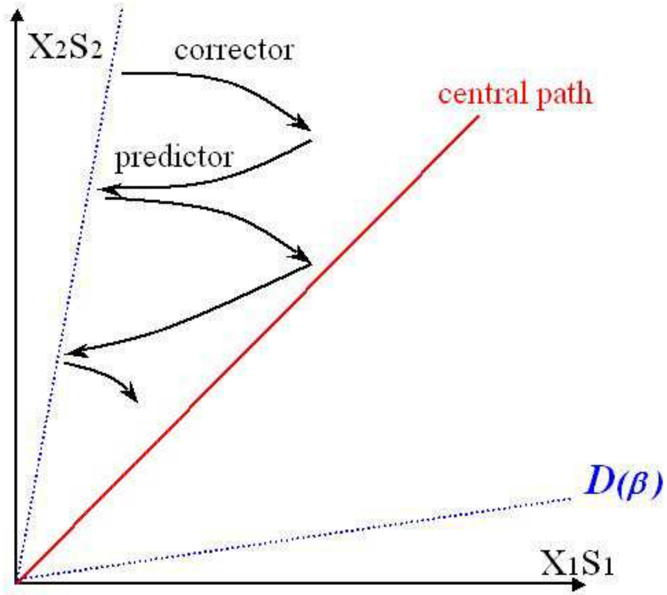
Remark If we take $\omega = 0.1$ then the values of $m = \lceil (n+2)^\omega - 1 \rceil$, corresponding to $n = 10^6$, $n = 10^7$, $n = 10^8$, and $n = 10^9$ are 3, 5, 6, and 7 respectively. This corresponds with efficient practical implementation of interior point methods where the same factorization is used from 3 to 7 times.

Classical Predictor-Corrector Methods



A classical predictor-corrector method operates in two neighborhoods of the central path. At each iteration, a predictor step is used to decrease the duality gap while keeping the point inside the outer neighborhood; a corrector step then follows to increase the centrality by bringing the iterated point back into the inner neighborhood.

Our Corrector-Predictor Method



Our corrector-predictor method (P., 2004) operates in one wide neighborhood $\mathcal{D}(\beta)$ of the central path. At each iteration, a corrector step is used to increase the centrality and optimality simultaneously. After the corrector step we obtain $\bar{z} \in \mathcal{D}(\beta_c)$, where $\beta_c > \beta$ varies from iteration to iteration. A predictor step then follows to increase the optimality while keeping the result point in $\mathcal{D}(\beta)$.

Since only one neighborhood is used we can extend to sufficient HLCP.

$P_*(\kappa)$ *and* P_* *HLCP*

- $P_*(\kappa)$ HLCP, $\kappa \geq 0$:

$$Qu + Rv = 0 \Rightarrow (1+4\kappa) \sum_{i\in\mathcal{I}^+} u_i v_i + \sum_{i\in\mathcal{I}^-} u_i v_i \geq 0, \text{ for any } u, v \in \mathbb{R}^n$$

  where $\mathcal{I}^+ = \{i : u_i v_i > 0\}, \quad \mathcal{I}^- = \{i : u_i v_i < 0\}, \forall u, v \in \mathbb{R}^n$.
  If the above condition is satisfied we say that (Q, R) is a $P_*(\kappa)$ pair.

- Monotone HLCP is a special case of $P_*(\kappa)$ HLCP with $\kappa = 0$.

- P_* HLCP:

$$P_* = \cup_{\kappa\geq 0} P_*(\kappa)\,,$$

  in which case (Q, R) is called a P_* pair.

Sufficient Matrices and Sufficient LCP

- A matrix $M \in R^{n \times n}$ is called
  - column sufficient, if $z_i(Mz)_i \leq 0$ for all $i \Rightarrow z_i(Mz)_i = 0$ for all i;
  - row sufficient, if its transpose is column sufficient;
  - sufficient, if it is both column and row sufficient.
- M is column sufficient iff for each $q \in R^n$, the $LCP(q, M)$ has a (possibly empty) convex solution set.
- M is row sufficient iff for each $q \in R^n$, if (x, u) is a KKT pair of the quadratic program

$$\begin{aligned} \min_x \quad & x^T(Mx + q) \\ \text{s.t.} \quad & Mx + q \geq 0 \\ & x \geq 0\,, \end{aligned}$$

  then x solves $LCP(q, M)$.
- (Väliaho, 1996) 'P_*-matrices are just sufficient'.

Sufficient Pairs and Sufficient HLCP

- $\Phi = Null([Q\ R])$. The pair (Q, R) is called
  - column sufficient, if $\forall \lceil u, v \rfloor \in \Phi$, $u_i v_i \leq 0 \Rightarrow u_i v_i = 0$;
  - row sufficient, if $\forall \lceil u, v \rfloor \in \Phi^{\perp}$, $u_i v_i \geq 0 \Rightarrow u_i v_i = 0$;
  - sufficient, if the pair is both column and row sufficient;

  where $\lceil u, v \rfloor$ denotes the vector $[u^T, v^T]^T$.
- For an HLCP, the following statements are equivalent:
  - It is a P_* HLCP;
  - It is a sufficient HLCP;
  - Its solution set is convex and every KKT point of

$$\begin{array}{ll} \min_{x,s} & x^T s \\ \text{s.t.} & Qx + Rs = b \\ & x, s \geq 0 \end{array}$$

  is a solution of the HLCP.

Interior Point Methods for $P_*(\kappa)$ and sufficient LCP

- Methods Using Small Neighborhoods
  - (Kojima et. al., 1991) Potential reduction algorithm for solving sufficient LCP, $O((1+\kappa)\sqrt{n}L)$ iteration complexity, no superlinear convergence.
  - (Miao, 1995) Generalization of MTY to $P_*(\kappa)$ LCP, $O((1+\kappa)\sqrt{n}L)$ iteration complexity and quadratic convergence (under the non-degeneracy assumption). The algorithm depends on κ.
  - (P. and Sheng, 1997) Generalization of MTY to sufficient LCP, $O((1+\kappa)\sqrt{n}L)$ iteration complexity and superlinear convergence (even in degenerate case). The algorithm does not depend on κ.

Interior Point Methods for $P_*(\kappa)$ and sufficient LCP

- Methods Using Wide Neighborhoods
  - (Stoer, Wechs, Mizuno, 1998) High order methods for sufficient HLCP. Superlinear convergence even in the degenerate case.
  - (Stoer, Wechs, 1998) $O((1+\kappa)\sqrt{n}L)$ iteration complexity in the small neighborhood of the central path.
  - (Stoer, 2001) Superlinear convergence in the large neighborhood of the central path. No complexity.
  - (Liu and P., 2005) A corrector-predictor method for sufficient HLCP in a wide neighborhood $\mathcal{D}(\beta)$ of the central path.
    - It has $O((1+\kappa)\sqrt{n}L)$ iteration complexity;
    - It is superlinearly convergent (even in the degenerate case);
    - It does not depend on κ, so that it can be used for sufficient HLCP.
    - The cost of implementing one iteration of our algorithm is $O(n^3)$ arithmetic operations.

Extensions

- Convex Programming (Nesterov and Nemirovskii)
  - self concordant barrier
  - universal barrier exists but no explicit expression
- LP and LCP over cones (Nesterov, Todd, ...)
- relation to Jordan algebras (Guler, Faybusovich)
- infinite dimensional problems (Faybusovich, Ulbrich, Heinkenschloss)
- "cone-free" (Nemirovskii and Tuncel)
- nonlinear programming (Nocedal, Vanderbei, Waechter **IPOPT**)

Semidefinite programming (SDP)

(Primal)

$$\begin{array}{ll} \text{minimize} & C \bullet X \\ \text{subject to} & A_i \bullet X = b_i, \; i = 1, \ldots, m, \; X \succeq 0. \end{array}$$

(Dual)

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \sum_{i=1}^{m} y_i A_i + S = C, \; S \succeq 0. \end{array}$$

data: $C, A_i,$ –$n \times n$ symmetric matrices; $b = (b_1, \ldots, b_m)^T \in \mathcal{R}^m$.
primal variable: X, symmetric & p.s.d.
dual variables: $y \in \mathcal{R}^m$, S, symmetric & p.s.d.

SDP central path

The primal-dual SDP system:

$$
\begin{aligned}
&A_i \bullet X = b_i, \; i = 1, \ldots, m, \\
&\sum_{i=1}^{m} y_i A_i + S = C, \\
&XS = 0, \quad X \succeq 0, \quad S \succeq 0.
\end{aligned}
$$

The primal-dual SDP central path:

$$
\begin{aligned}
&A_i \bullet X = b_i, \; i = 1, \ldots, m, \\
&\sum_{i=1}^{m} y_i A_i + S = C, \\
&XS = \mu I, \quad X \succeq 0, \quad S \succeq 0.
\end{aligned}
$$

Problem: On the central path $XS = SX$, but this is not true outside the central path.

Search directions

MZ search direction $(\Delta X, \Delta y, \Delta S)$:

$$H_P(X\Delta S + \Delta X S) = \sigma\mu I - H_P(XS),$$

$$A_i \bullet \Delta X = r_i, \quad i = 1, \ldots, m,$$

$$\sum_{i=1}^{m} \Delta y_i A_i + \Delta S = R_d.$$

Symmetrization operator:

$$H_P(M) = (PMP^{-1} + [PMP^{-1}]^T)/2.$$

$P = I$: AHO
$P = S^{1/2}$: HKM (HRVW/KSH/M)
P such that $P^T P = X^{-1/2}[X^{1/2} S X^{1/2}]^{1/2} X^{-1/2}$: NT

Twenty search directions are analyzed and tested by Todd (1999)

MTY with some directions has $O(\sqrt{n}\ln(\epsilon_0/\epsilon))$ iteration complexity.

Superlinear Convergence for SDP

Kojima, Shida and Shindoh (1998): MTY predictor-corrector with HKM search direction has superlinear convergence if:

(A) SDP has a strictly complementary solution;

(B) SDP is nondegenerate (nonsingular Jacobian)

(C) the iterates converge tangentially to the central path in the sense that the size of the neighborhood containing the iterates must approach zero, namely,

$$\frac{\left\|(X^k)^{1/2} S^k (X^k)^{1/2} - (X^k \bullet S^k/n) I\right\|_F}{X^k \bullet S^k} \to 0$$

Assumption (B) and (C) not required for LP.

Superlinear Convergence for SDP – continued

P. and Sheng (1998): Superlinear convergence + polynomiality if

(A) SDP has a strictly complementary solution;

(D)
$$\frac{X^k S^k}{\sqrt{X^k \bullet S^k}} \to 0.$$

- Note that (B) and (C) implies (D).
- Both (C) and (D) can be enforced by the algorithm; practical efficiency of such an approach is questionable. I
- f several corrector steps are used the algorithm has polynomial complexity and is superlinearly convergent under assumption (A) only.
- MTY with KHM direction for predictor and AHO direction for corrector, has polynomial complexity and superlinear convergence of Q-order 1.5 under (A) and (B).

Superlinear Convergence for SDP – wanted

Kojima, Shida and Shindoh (1998):

- example suggesting that interior point algorithms for SDP based on the KHM are unlikely to be superlinearly convergent without(C).
- MTY with AHO is quadratically convergent under (A). Global convergence, but no polynomial complexity.

Ji, P. and Sheng (1999): MTY using the MZ-family.

- polynomial complexity (Monteiro).
- (A) + (D)$\Rightarrow$ superlinear convergence.
- (A) + (B) + scaling matrices in the corrector step have bounded condition number $\Rightarrow$ Q-order 1.5 .
- (A) + (B) + scaling matrices in both predictor and corrector step have bounded condition number $\Rightarrow$ Q-quadratic convergence.

Superlinear Convergence for SDP – wanted

Kojima, Shida and Shindoh (1998):

- example suggesting that interior point algorithms for SDP based on the KHM are unlikely to be superlinearly convergent without(C).
- MTY with AHO is quadratically convergent under (A). Global convergence, but no polynomial complexity.

Ji, P. and Sheng (1999): MTY using the MZ-family.

- polynomial complexity (Monteiro).
- (A) + (D)$\Rightarrow$ superlinear convergence.
- (A) + (B) + scaling matrices in the corrector step have bounded condition number $\Rightarrow$ Q-order 1.5 .
- (A) + (B) + scaling matrices in both predictor and corrector step have bounded condition number $\Rightarrow$ Q-quadratic convergence.

WANTED: A superlinearly convergent algorithm with $O(\sqrt{n}\ln(\epsilon_0/\epsilon))$ iteration complexity in a wide neighborhood of the central path.

References

Detailed expositions of the results obtained in interior point methods prior to 1997 are contained in the monographs of Roos, Vial and Terlaky (1997), Wright (1997), and Ye (1997). For a short survey of the results obtained prior to 2000 see P. and Wright (2000). In what follows we give complete references for the results cited in this talk:

- Anst99 K.M. Anstreicher. Linear programming in $O([n^3/\ln\, n]L)$ operations. CORE Discussion Paper 9746, Universite Catholique de Louvain, Louvain-la-Neuve, Belgium, January 1999. to appear in SIAM Journal on Optimization.
- BoGo96 J. F. Bonnans and C. C. Gonzaga. Convergence of interior point algorithms for the monotone linear complementarity problem. *Mathematics of Operations Research*, 21:1–25, 1996.
- BoPo97 J. F. Bonnans and F. A. Potra. On the convergence of the iteration sequence of infeasible path following algorithms for linear complementarity problems. *Mathematics of Operations Research*, 22(2):378–407, 1997.

References–continued

- FrMi96 R. M. Freund and S. Mizuno. Interior point methods: Current status and future directions. *Optima*, 51:1–9, October 1996.
- Gonz99 C. C. Gonzaga. Complexity of predictor-corrector algorithms for LCP based on a large neighborhood of the central path. *SIAM J. Optim.*, 10(1):183–194 (electronic), 1999.
- GoTa97a C. C. Gonzaga and R. A. Tapia. On the convergence of the Mizuno-Todd-Ye algorithm to the analytic center of the solution set. *SIAM Journal on Optimization*, 7(1):47–65, 1997.
- GoTa97b C. C. Gonzaga and R. A. Tapia. On the quadratic convergence of the simplified Mizuno-Todd-Ye algorithm for linear programming. *SIAM Journal on Optimization*, 7(1):66–85, 1997.
- JiPS99 J. Ji, F. A. Potra, and R. Sheng. On the local convergence of a predictor-corrector method for semidefinite programming. *SIAM Journal on Optimization*, 10:195–210, 1999.
- Karm84 N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

References–continued

- Khac79 L. G. Khachiyan. A polynomial algorithm for linear programming. *Soviet Mathematics Doklady*, 20:191–194, 1979.
- KlMi72 V. Klee and G. Minty. How good is the simplex algorithm? In O. Sisha, editor, *Inequalities III*. Academic Press, New York, NY, 1972.
- KoSS98 Masakazu Kojima, Masayuki Shida, and Susumu Shindoh. Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. *Math. Programming*, 80(2, Ser. A):129–160, 1998.
- MeYe93 S. Mehrotra and Y. Ye. Finding an interior point in the optimal face of linear programs. *Math. Programming*, 62(3, Ser. A):497–515, 1993.
- Potr01 F. A. Potra. Q-superlinear convergence of the iterates in primal-dual interior-point methods. *Math. Programming*, 91(Ser. A):99–115, 2001.

References–continued

- Potr03 F. A. Potra. A superlinearly convergent predictor-corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{n}l)$-iteration complexity. *Math. Programming*, (Ser. A, electronic):DOI 10.1007/s10107=003–0472–0, 2003.
- PoSg98b F. A. Potra and R. Sheng. Superlinear convergence of interior-point algorithms for semidefinite programming. *Journal of Optimization Theory and Applications*, 99(1):103–119, 1998.
- PoSg98c F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. *SIAM Journal on Optimization*, 8(4):1007–1028, 1998.
- PoWr00 F. A. Potra and S. J. Wright. Interior-point methods. *Journal of Computational and Applied Mathematics*, 124:281–302, 2000.
- RoVT97 C. Roos, J.-Ph. Vial, and T. Terlaky. *Theory and Algorithms for Linear Optimization : An Interior Point Approach*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.

References–continued

- Sham92 R. Shamir. Probabilistic analysis in linear programming.
- TaZY95 R. A. Tapia, Y. Zhang, and Y. Ye. On the convergence of the iteration sequence in primal–dual interior-point methods. *Mathematical Programming*, 68(2):141–154, 1995.
- Todd99 M. J. Todd. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Optim. Methods Softw.*, 11/12(1-4):1–46, 1999. Interior point methods.
- Wrig97 S. J. Wright. *Primal–Dual Interior–Point Methods*. SIAM Publications, Philadephia, 1997.
- Ye97 Y. Ye. *Interior Point Algorithms : Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.
- YeAn93 Y. Ye and K. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP. *Mathematical Programming*, 62(3):537–551, 1993.

References–continued

- YGTZ93 Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. *Mathematical Programming*, 59(2):151–162, 1993.
- ZhTD92 Y. Zhang, R. A. Tapia, and J. E. Dennis. On the superlinear and quadratic convergence of primal–dual interior point linear programming algorithms. *SIAM Journal on Optimization*, 2(2):304–324, 1992.
- ZhTP93 Y. Zhang, R. A. Tapia, and F. A. Potra. On the superlinear convergence of interior point algorithms for a general class of problems. *SIAM J. on Optimization*, 3(2):413–422, 1993.