Numerical Optimization

Lecture 7: Interior Point Method

王浩

信息科学与技术学院

Email: wanghao1@shanghaitech.edu.cn

Ellipsoid Method
（椭球法）

Some Milestones in the History of LO

- 1947 simplex method – Dantzig; still very efficient
- 1972 exponential example – Klee–Minty – theoretical $O(2^n)$
- 1979 ellipsoid method $O(n^6L^2)$ – Leonid Khachiyan (Леонид Генрихович Хачиян); extended to QP
- 1979–1985 ellipsoid method and combinatorial optimization
- 1984 projective IPM $O(n^{3.5}L^2)$ – Karmarkar – efficient in practice!?
- ....

Standard form LO: Fundamentals

$$(P) \qquad \min\{c^T x : Ax = b, x \geq 0\},$$
$$(D) \qquad \max\{b^T y : A^T y + s = c, s \geq 0\}$$

A is an $m \times n$ matrix with rand$(A) = m$

Proposition 1 (Weak duality). If $x \in \mathbb{R}^n$ is primal, $y \in \mathbb{R}^m$ is dual feasible, then $c^T x \geq b^T y$, where the $c^T x = b^T y$ is satisfied iff $x^T s = 0$.

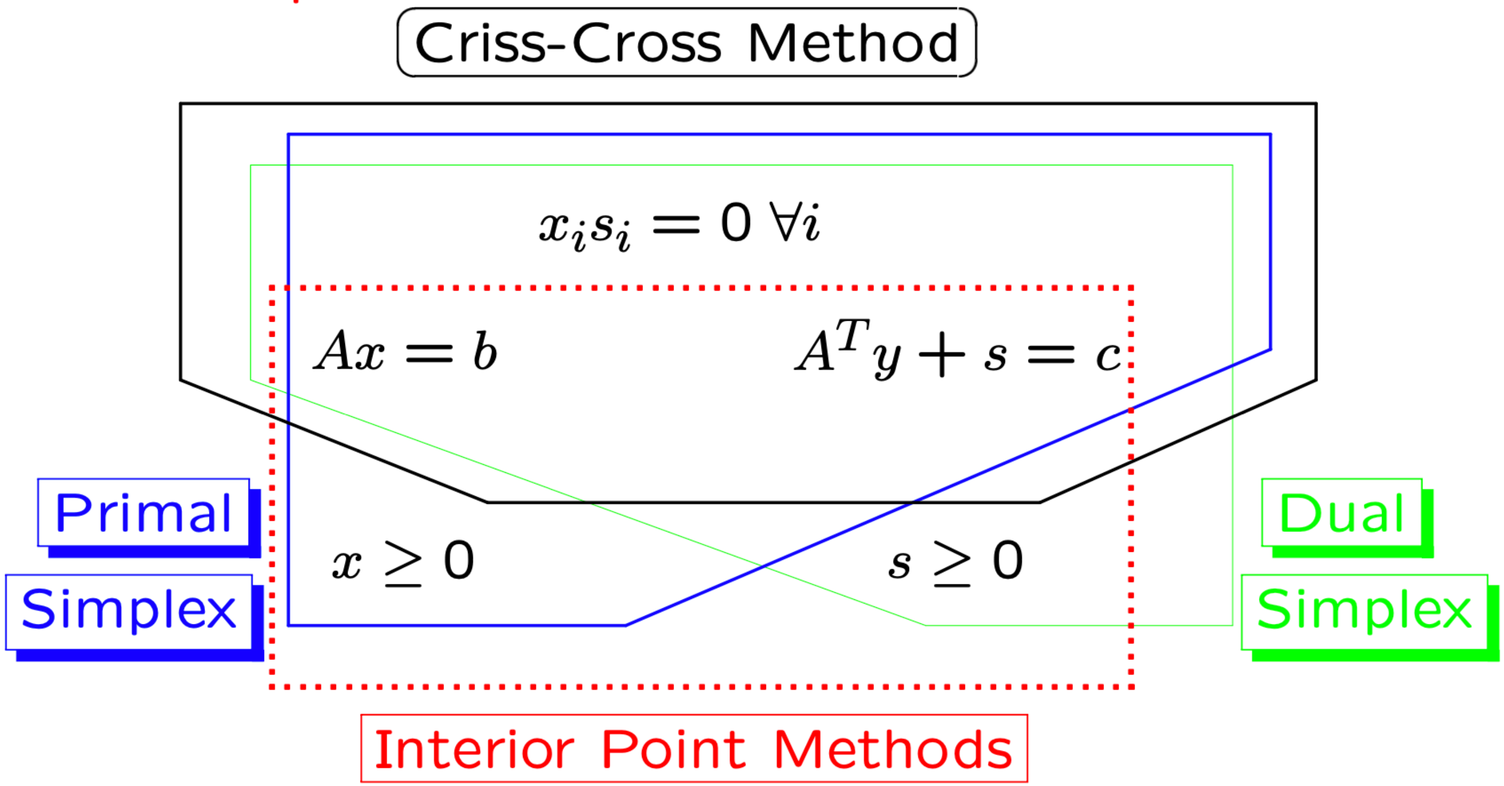
Corollary 1 (Complementarity). If $x \in \mathbb{R}^n$ is primal, $y \in \mathbb{R}^m$ is dual feasible and $x^T s = 0$, then x and y are primal and dual optimal respectively.

Optimality: optimal solutions can be given as the set of solutions of the system:

$$Ax = b, x \geq 0 \qquad\qquad A^T y + s = c, s \geq 0$$
$$x_i s_i = 0, \ \forall i$$

Linear Optimization: Fundamentals

Algorithmic concepts



Ellipsoid method: $c^T x \leq b^T y$; **None is satisfied**

Khachiyan and the NYT

The New York Times

NEW YORK, WEDNESDAY, NOVEMBER 7, 1979

A Soviet Discovery Rocks World of Mathematics

By MALCOLM W. BROWNE

A surprise discovery by an obscure Soviet mathematician has rocked the world of mathematics and computer analysis, and experts have begun exploring its practical applications.

Mathematicians describe the discovery by L.G. Khachian as a method by which computers can find guaranteed solutions to a class of very difficult problems that have hitherto been tackled on a kind of hit-or-miss basis.

Apart from its profound theoretical interest, the discovery may be applicable in weather prediction, complicated industrial processes, petroleum refining, the scheduling of workers at large factories, secret codes and many other things.

"I have been deluged with calls from virtually every department of government for an interpretation of the significance of this," a leading expert on computer methods, Dr. George B. Dantzig of Stanford University, said in an interview.

The solution of mathematical problems by computer must be broken down into a series of steps. One class of problem sometimes involves so many steps that it could take billions of years to compute.

The Russian discovery offers a way by which the number of steps in a solution can be dramatically reduced. It also offers the mathematician a way of learning quickly whether a problem has a solution or not, without having to complete the entire immense computation that may be required.

According to the American journal Sci-

Continued on Page A20, Column 3

Russian's Surprise Problem-Solving Discovery Reported

Continued From Page A1

TUESDAY, NOVEMBER 27, 1979

Science Times

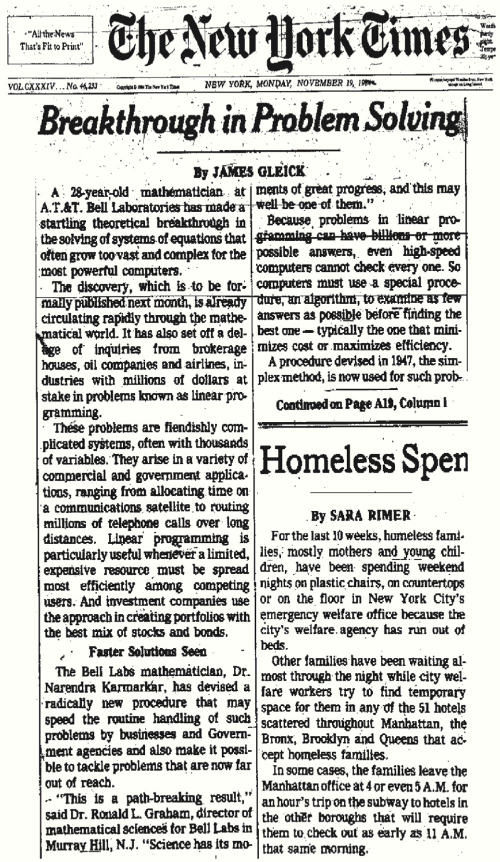
The New York Times

Soviet Mathematician Is Obscure No More

Mathematician Is Obscure No More

An Approach to Difficult Problems

Karmarkar and the NYT



"All the News That's Fit to Print"

The New York Times

VOL. CXXXIV . . . No. 46,233 — NEW YORK, MONDAY, NOVEMBER 19, 1984

Breakthrough in Problem Solving

By JAMES GLEICK

A 28-year-old mathematician at A.T.&T. Bell Laboratories has made a startling theoretical breakthrough in the solving of systems of equations that often grow too vast and complex for the most powerful computers.

The discovery, which is to be formally published next month, is already circulating rapidly through the mathematical world. It has also set off a deluge of inquiries from brokerage houses, oil companies and airlines, industries with millions of dollars at stake in problems known as linear programming.

These problems are fiendishly complicated systems, often with thousands of variables. They arise in a variety of commercial and government applications, ranging from allocating time on a communications satellite to routing millions of telephone calls over long distances. Linear programming is particularly useful whenever a limited, expensive resource must be spread most efficiently among competing users. And investment companies use the approach in creating portfolios with the best mix of stocks and bonds.

Faster Solutions Seen

The Bell Labs mathematician, Dr. Narendra Karmarkar, has devised a radically new procedure that may speed the routine handling of such problems by businesses and Government agencies and also make it possible to tackle problems that are now far out of reach.

"This is a path-breaking result," said Dr. Ronald L. Graham, director of mathematical sciences for Bell Labs in Murray Hill, N.J. "Science has its moments of great progress, and this may well be one of them."

Because problems in linear programming can have billions or more possible answers, even high-speed computers cannot check every one. So computers must use a special procedure, an algorithm, to examine as few answers as possible before finding the best one — typically the one that minimizes cost or maximizes efficiency.

A procedure devised in 1947, the simplex method, is now used for such prob-

Continued on Page A19, Column 1

Homeless Spen

By SARA RIMER

For the last 10 weeks, homeless families, mostly mothers and young children, have been spending weekend nights on plastic chairs, on countertops or on the floor in New York City's emergency welfare office because the city's welfare agency has run out of beds.

Other families have been waiting almost through the night while city welfare workers try to find temporary space for them in any of the 51 hotels scattered throughout Manhattan, the Bronx, Brooklyn and Queens that accept homeless families.

In some cases, the families leave the Manhattan office at 4 or even 5 A.M. for an hour's trip on the subway to hotels in the other boroughs that will require them to check out as early as 11 A.M. that same morning.

AT&T Announces the KORBX System

NEWS CLIP

AT&T Markets Problem Solver, Based On Math Whiz's Find, for $8.9 Million

By ROGER LOWENSTEIN
Staff Reporter of THE WALL STREET JOURNAL

NEW YORK—American Telephone & Telegraph Co. has called its math whiz, Narendra Karmarkar, a latter-day Isaac Newton. Now, it will see if he can make the firm some money.

Four years after AT&T announced an "astonishing" discovery by the Indian-born Mr. Karmarkar, it is marketing an $8.9 million problem solver based on his invention.

Dubbed Korbx, the computer-based system is designed to solve major operational problems of both business and government. AT&T predicts "substantial" sales for the product, but outsiders say the price is high and point out that its commercial viability is unproven.

"At $9 million a system, you're going to have a small number of users," says Thomas Magnanti, an operations-research specialist at Massachusetts Institute of Technology. "But for very large-scale problems, it might make the difference."

Korbx uses a unique algorithm, or step-by-step procedure, invented by Mr. Karmarkar, a 32-year old, an AT&T Bell Laboratories mathematician.

"It's designed to solve extremely difficult or previously unsolvable resource-allocation problems—which can involve hundreds of thousands of variables—such as personnel planning, vendor selection, and equipment scheduling," says Aristides Fronistas, president of an AT&T division created to market Korbx.

Potential customers might include an airline trying to determine how to route many planes between numerous cities and an oil company figuring how to feed different grades of crude oil into various refineries and have the best blend of refined products emerge.

AT&T says that fewer than 10 companies, which it won't name, are already using Korbx. It adds that, because of the price, it is targeting only very large companies—mostly in the Fortune 100.

Korbx "won't have a significant bottom-line impact initially" for AT&T, though it might in the long term, says Charles Nichols, an analyst with Bear, Stearns & Co. "They will have to expose it to users and demonstrate" it uses.

AMR Corp.'s American Airlines says it's considering buying AT&T's system. Like other airlines, the Fort Worth, Texas, carrier has the complex task of scheduling pilots, crews and flight attendants on thousands of flights every month.

Thomas M.Cook, head of operations research at American, says, "Every airline has programs that do this. The question is: Can AT&T do it better and faster? The jury is still out."

The U.S. Air Force says it is considering using the system at the Scott Air Force Base in Illinois.

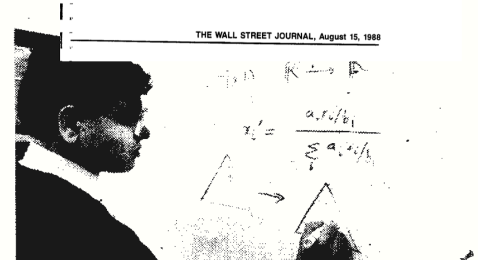
One reason for the uncertainty is that AT&T has, for reasons of commercial secrecy, deliberately kept the specifics of Mr. Karmarkar's algorithm under wraps.

"I don't know the details of their system," says Eugene Bryan, president of Decision Dynamics Inc., a Portland, Ore., consulting firm that specializes in linear programming, a mathematical technique that employs a series of equations using many variables to find the most efficient way of allocating resources.

Mr. Bryan says, though, that if the Karmarkar system works, it would be extremely useful. "For every dollar you spend on optimization," he says, "you usually get them back many-fold."

AT&T has used the system in-house to help design equipment and routes on its Pacific Basin system, which involves 22 countries. It's also being used to plan AT&T's evolving domestic network, a problem involving some 800,000 variables.

THE WALL STREET JOURNAL, August 15, 1988



The New York Times/Keith Meyers

Dr. Narendra Karmarkar working on his algorithm at Bell Labs in Murray Hill, N.J. His method appears to be much faster than current approaches.

Theoretical Breakthrough Offers New Insights to Problem Solving

Breakthrough in Problem Solving

By JAMES GLEICK

A 28-year-old mathematician at A.T.&T. Bell Laboratories has made a startling theoretical breakthrough in the solving of systems of equations that often grow too vast and complex for the most powerful computers.

The discovery, which is to be formally published next month, is already circulating rapidly through the mathematical world. It has also set off a deluge of inquiries from brokerage houses, oil companies and airlines, industries with millions of dollars at stake in problems known as linear programming.

Faster Solutions Seen

These problems are fiendishly complicated systems, often with thousands of variables. They arise in a variety of commercial and government applications, ranging from allocating time on a communications satellite to routing millions of telephone calls over long distances, or whenever a limited, expensive resource must be spread most efficiently among competing users. And investment companies use them in creating portfolios with the best mix of stocks and bonds.

The Bell Labs mathematician, Dr. Narendra Karmarkar, has devised a radically new procedure that may speed the routine handling of such problems by businesses and Government agencies and also make it possible to tackle problems that are now far out of reach.

"This is a path-breaking result," said Dr. Ronald L. Graham, director of mathematical sciences for Bell Labs in Murray Hill, N.J.

"Science has its moments of great progress, and this may well be one of them."

Because problems in linear programming can have billions or more possible answers, even high-speed computers cannot check every one. So computers must use a special procedure, an algorithm, to examine as few answers as possible before finding the best one — typically the one that minimizes cost or maximizes efficiency.

A procedure devised in 1947, the simplex method, is now used for such problems,

Continued on Page A19, Column 1

《纽约时报》

THE NEW YORK TIMES, November 19, 1984

Karmarkar at Bell Labs: an equation to find a new way through the maze

Folding the Perfect Corner

A young Bell scientist makes a major math breakthrough

Every day 1,200 American Airlines jets crisscross the U.S., Mexico, Canada and the Caribbean, stopping in 110 cities and bearing over 80,000 passengers. More than 4,000 pilots, copilots, flight personnel, maintenance workers and baggage carriers are shuffled among the flights; a total of 3.6 million gal. of high-octane fuel is burned. Nuts, bolts, altimeters, landing gears and the like must be checked at each destination. And while performing these scheduling gymnastics, the company must keep a close eye on costs, projected revenue and profits.

Like American Airlines, thousands of companies must routinely untangle the myriad variables that complicate the efficient distribution of their resources. Solving such monstrous problems requires the use of an abstruse branch of mathematics known as linear programming. It is the kind of math that has frustrated theoreticians for years, and even the fastest and most powerful computers have had great difficulty juggling the bits and pieces of data. Now Narendra Karmarkar, a 28-year-old Indian-born mathematician at Bell Laboratories in Murray Hill, N.J., after only a years' work has cracked the puzzle of linear programming by devising a new algorithm, a step-by-step mathematical formula. He has translated the procedure into a program that should allow computers to track a greater combination of tasks than ever before and in a fraction of the time.

Unlike most advances in theoretical mathematics, Karmarkar's work will have an immediate and major impact on the real world. "Breakthrough is one of the most abused words in science," says Ronald Graham, director of mathematical sciences at Bell Labs. "But this is one situation where it is truly appropriate."

Before the Karmarkar method, linear equations could be solved only in a cumbersome fashion, ironically known as the simplex method, devised by Mathematician George Dantzig in 1947. Problems are conceived of as giant geodesic domes with thousands of sides. Each corner of a facet on the dome

《时代周刊》

TIME MAGAZINE, December 3, 1984

Karmarkar Algorithm Proves Its Worth

Less than two years after discovery of a mathematical procedure that Bell Labs said could solve a broad range of complex business problems 50 to 100 times faster than current methods, AT&T is filing for patents covering its use. The Karmarkar algorithm, which drew headlines when discovered by researcher Narendra Karmarkar, will be applied first to AT&T's long-distance network.

Thus far, Bell Labs has verified the procedure's capabilities in developing plans for new fiber-optic transmission and satellite capacity linking 20 countries bordering the Pacific Ocean. That jointly owned network will be built during the next 10 years. Planning requires a tremendous number of "what if" scenarios involving 43,000 variables describing transmission capacity, location and construction schedules, all juggled amid political considerations of each connected country.

The Karmarkar algorithm was able to solve the Pacific Basin problem in four minutes, against 80 minutes by the method previously used, says Neil Dinn, head of Bell Labs' international transmission planning department. The speedier solutions will enable international committees to agree on network designs at one meeting instead of many meetings stretched out over months.

AT&T now is using the Karmarkar procedure to plan construction for its domestic network, a problem involving 800,000 variables. In addition, the procedure may be written into software controlling routing of domestic phone calls, boosting the capacity of AT&T's current network.

《华尔街日报》

THE WALL STREET JOURNAL, July 18, 1986

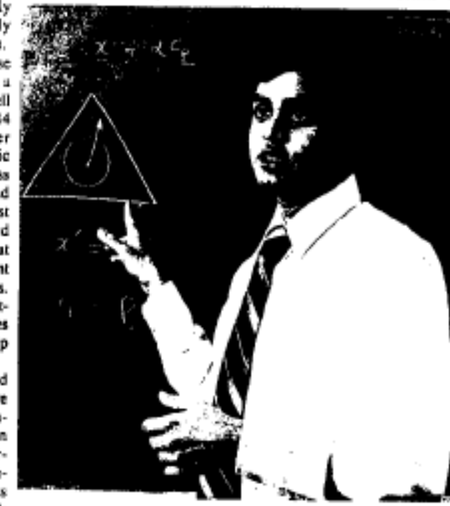
THE STARTLING DISCOVERY BELL LABS KEPT IN THE SHADOWS

Now its breakthrough mathematical formula could save business millions

It happens all too often in science. An obscure researcher announces a stunning breakthrough and achieves instant fame. But when other scientists try to repeat his results, they fail. Fame quickly turns to notoriety, and eventually the episode is all but forgotten.

That seemed to be the case with Narendra K. Karmarkar, a young scientist at AT&T Bell Laboratories. In late 1984 the 28-year-old researcher astounded not only the scientific community but also the business world. He claimed he had cracked one of the thorniest aspects of computer-aided problem-solving. If so, his feat would have meant an instant windfall for many big companies. It could also have pointed to better software for small companies that use computers to help manage their business.

Karmarkar said he had discovered a quick way to solve problems so hideously complicated that they often defy even the most powerful supercomputers. Such problems bedevil a broad range of business activities, from assessing risk factors in stock portfolios to drawing up production schedules in factories. Just about any company that distributes products through more than a handful of warehouses bumps into such problems when calculating the cheapest routes for getting goods to customers. Even when the problems aren't terribly complex, solving them can chew up so much computer time that the answer is useless before it's found.



KARMARKAR: SKEPTICS ATTACKED HIS PRECOCIOUS FEAT

HEAD START. To most mathematicians, Karmarkar's precocious feat was hard to swallow. Because such questions are so common, a special branch of mathematics called linear programming (LP) has evolved, and most scientists thought that was as far as they could go. Sure enough, when other researchers independently tried to test Karmarkar's process, their results were disappointing. At scientific conferences skeptics attacked the algorithm's validity as well as Karmarkar's veracity.

But this story may end with a different twist. Other scientists weren't able to duplicate Karmarkar's work, it turns out, because his employer wanted it that way. Vital details about how best to translate the algorithm, whose mathematical notations run on for about 20 printed pages, into digital computer code were withheld to give Bell Labs a head start at developing commercial products. Following the breakup of American Telephone & Telegraph Co. in January, 1984, Bell Labs was no longer prevented from exploiting its research for profit. While the underlying concept could not be patented or copyrighted because it is pure knowledge, any computer programs that AT&T developed to implement the procedure can be protected.

Now, AT&T may soon be selling the first product based on Karmarkar's work—to the U.S. Air Force. It includes a multiprocessor computer from Alliant Computer Systems Corp. and a software version of Karmarkar's algorithm that has been optimized for high-speed parallel processing. The system would be installed at St. Louis' Scott Air Force Base, headquarters of the Military Airlift Command (MAC). Neither party will comment on the deal's cost or where the negotiations stand, but the Air Force's interest is easy to fathom.

JUGGLING ACT. On a typical day thousands of planes ferry cargo and passengers among air fields scattered around the world. To keep those jets flying, MAC

BUSINESS WEEK, September 21, 1987

《商业周刊》

Patents

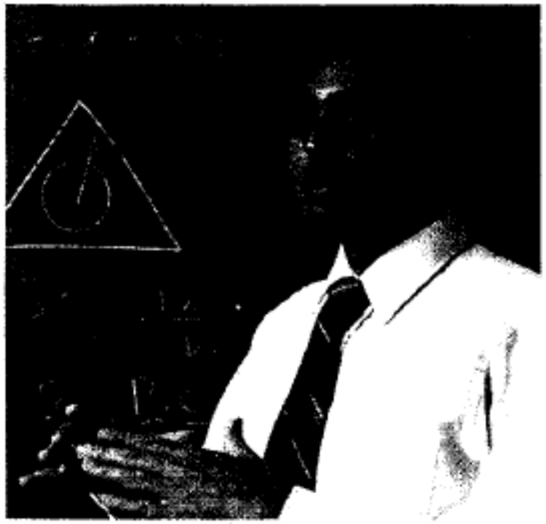
by Stacy V. Jones

A Method to Improve Resource Allocation

Scientists at Bell Laboratories in Murray Hill, N.J., were granted three patents this week for methods of improving the efficiency of allocation of industrial and commercial resources.

The American Telephone and Telegraph Company, the laboratory's sponsor, is using the methods internally to regulate such operations as long-distance services.

Narendra K. Karmarkar of the laboratory staff was granted patent 4,744,028 for methods of allocating telecommunication and other resources. With David A. Bayer and Jeffrey C. Lagarian as co-inventors, he was granted patent 4,744,027 on improvements of the basic method. Patent 4,744,026 went to Robert J. Vanderbei for enhanced procedures.



Narendra K. Karmarkar of the Bell Laboratories staff.

THE NEW YORK TIMES, May 14, 1988

《纽约时报》

AT&T Markets Problem Solver, Based On Math Whiz's Find, for $8.9 Million

By ROGER LOWENSTEIN

Staff Reporter of THE WALL STREET JOURNAL

NEW YORK—American Telephone & Telegraph Co. has called its math whiz, Narendra Karmarkar, a latter-day Isaac Newton. Now, it will see if he can make the firm some money.

Four years after AT&T announced an "astonishing" discovery by the Indian-born Mr. Karmarkar, it is marketing an $8.9 million problem solver based on his invention.

Dubbed Korbx, the computer-based system is designed to solve major operational problems of both business and government. AT&T predicts "substantial" sales for the product, but outsiders say the price is high and point out that its commercial viability is unproven.

"At $9 million a system, you're going to have a small number of users," says Thomas Magnanti, an operations-research specialist at Massachusetts Institute of Technology. "But for very large-scale problems, it might make the difference."

Korbx uses a unique algorithm, or step-by-step procedure, invented by Mr. Karmarkar, a 32-year old, an AT&T Bell Laboratories mathematician.

"It's designed to solve extremely difficult or previously unsolvable resource-allocation problems—which can involve hundreds of thousands of variables—such as personnel planning, vendor selection, and equipment scheduling," says Aristides Fronistas, president of an AT&T division created to market Korbx.

Potential customers might include an airline trying to determine how to route many planes between numerous cities and an oil company figuring how to feed different grades of crude oil into various refineries and have the best blend of refined products emerge.

AT&T says that fewer than 10 companies, which it won't name, are already using Korbx. It adds that, because of the price, it is targeting only very large companies—mostly in the Fortune 100.

Korbx "won't have a significant bottom-line impact initially" for AT&T, though it might in the long term, says Charles Nichols, an analyst with Bear, Stearns & Co. "They will have to expose it to users and demonstrate" it uses.

AMR Corp.'s American Airlines says it's considering buying AT&T's system. Like other airlines, the Fort Worth, Texas, carrier has the complex task of scheduling pilots, crews and flight attendants on thousands of flights every month.

Thomas M.Cook, head of operations research at American, says, "Every airline has programs that do this. The question is: Can AT&T do it better and faster? The jury is still out."

The U.S. Air Force says it is considering using the system at the Scott Air Force Base in Illinois.

One reason for the uncertainty is that AT&T has, for reasons of commercial secrecy, deliberately kept the specifics of Mr. Karmarkar's algorithm under wraps.

"I don't know the details of their system," says Eugene Bryan, president of Decision Dynamics Inc., a Portland, Ore., consulting firm that specializes in linear programming, a mathematical technique that employs a series of equations using many variables to find the most efficient way of allocating resources.

Mr. Bryan says, though, that if the Karmarkar system works, it would be extremely useful. "For every dollar you spend on optimization," he says, "you usually get them back many-fold."

AT&T has used the system in-house to help design equipment and routes on its Pacific Basin system, which involves 22 countries. It's also being used to plan AT&T's evolving domestic network, a problem involving some 800,000 variables.

《华尔街日报》

THE WALL STREET JOURNAL, August 15, 1988

Polynomial Complexity and Condition Numbers

Let $f(n)$ and $g(n)$ be positive functions on the natural numbers. Then

- $f(n) = O(g(n))$ if $\exists c, n_0$, s.t. $f(n) \leq cg(n)$ for all $n \geq n_0$
- $f(n) = \Omega(g(n))$ if $\exists c, n_0$, s.t. $f(n) \geq cg(n)$ for all $n \geq n_0$
- $f(n) = \Theta(g(n))$ if $\exists c, n_0$, s.t. $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$

Solve $A^T y \leq c$ with integer data, $A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n$

Binary Input Length:

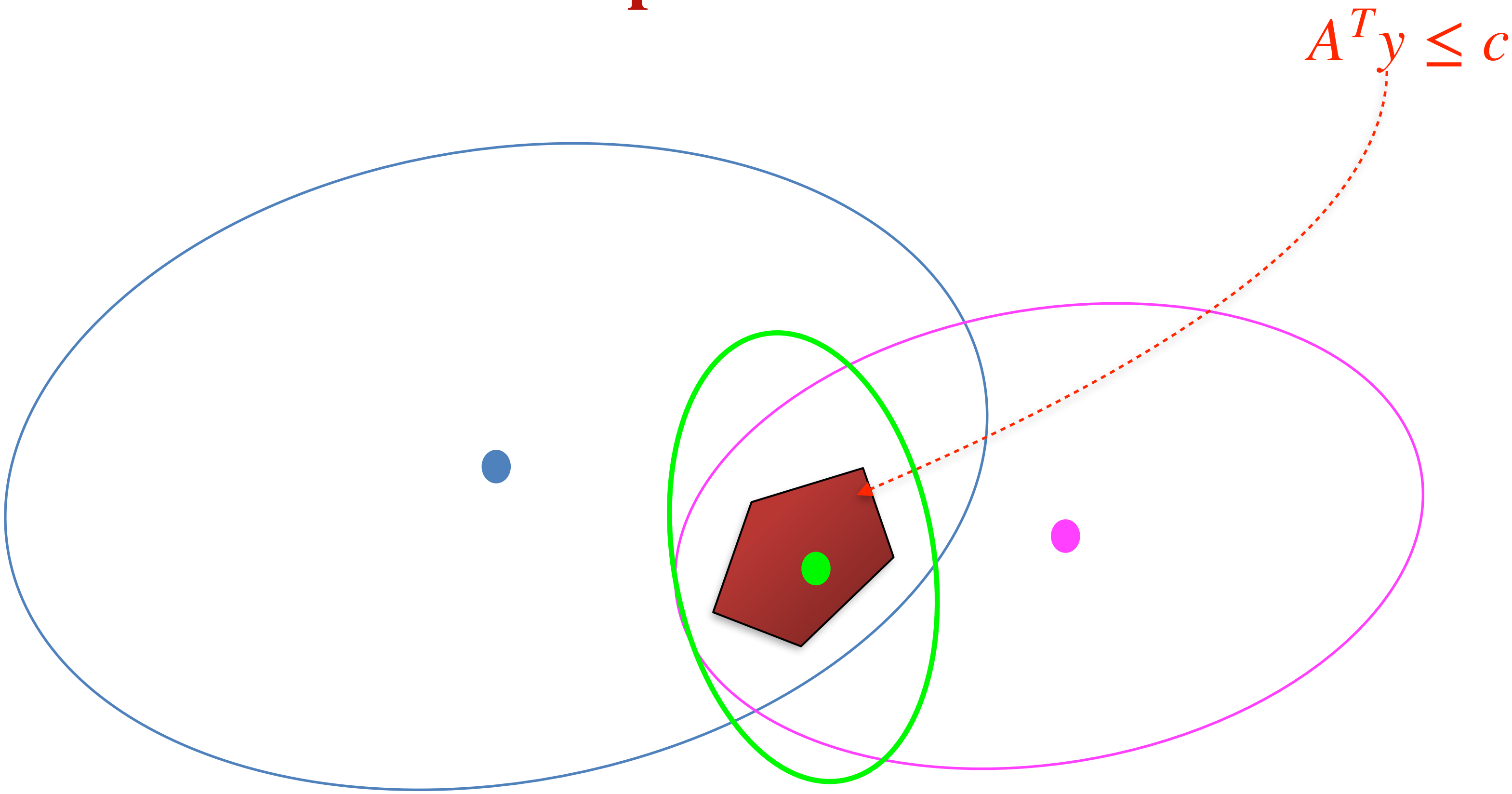
$$L = mn + n + \sum_{i=1}^{n} \sum_{j=1}^{m} \lceil \log_2(|a_{ij}| + 1) \rceil + \sum_{i=1}^{n} \lceil \log_2(|c_j| + 1) \rceil$$

Size of data:

Encoding an integer n in binary takes $\lceil \log_2(n+1) \rceil + 1$ bits

$$\sigma = \lceil \|c\| \Pi_{i=1}^{n} \|a_i\| \rceil$$

Basic idea of the Ellipsoid Method



Fundamentals of the Ellipsoid Method

Equivalent feasible sets: Let $\sigma = \|c\| \cdot \|a_1\| \cdot \ldots \cdot \|a_n\|$. Inequality system $A^T y \le c$ is consistent if and only if the inequality system

$$A^T y \le c + \frac{1}{mn\sigma} e, \qquad |y_i| \le \sigma + \frac{1}{mn\sigma} \text{ for all } i$$

is consistent, that contains a cube of size at least $\dfrac{1}{m^2 n \sigma}$.

Shrinking ellipsoids:

$P: m \times m$ positive semidefinite, $z \in \mathbb{R}^m$ gives the ellipsoid

$$\mathscr{E} = \text{ell} = \{y \in \mathbb{R}^m \mid (y-z)^T P^{-1} (y-z)\}.$$

Given $\mathscr{E} = \text{ell}(z, P)$ and $a \in \mathbb{R}^m$. Let $\mathscr{E}' = \text{ell}(z', P')$ be given by

$$z' = z - \frac{1}{m+1} \frac{Pa}{\sqrt{a^T P a}}, \quad P' = \frac{m^2}{m^2 - 1} \left\{ P - \frac{2}{m+1} \frac{P a a^T P}{a^T P a} \right\}$$

Fact: Ellipsoid $\mathscr{E}'$ contents the half ellipsoid $E \cap \{x \mid a^T (x - z) \le 0\}$

and $\text{vol}(\mathscr{E}')/\text{vol}(\mathscr{E}) < \exp\left(\dfrac{-1}{2(m+1)}\right)$

The Ellipsoid Method

Input:

Given $\bar{A}^T y \leq \bar{c}$ with at least $1/m^2 n\sigma$ size cube inside;

$\sigma = \|\bar{c}\| \cdot \|\bar{a}_1\| \cdot \ldots \cdot \|\bar{a}_n\|$; let $z^0 = 0$; $\quad P = 2\sigma I$, $\quad k = 0$.

begin

If $\bar{A}^T z^k \leq \bar{c}$ then STOP.

Else find i s.t. $\bar{a}_i^T z^k > \bar{c}_i$

Let $a = a_i$ and

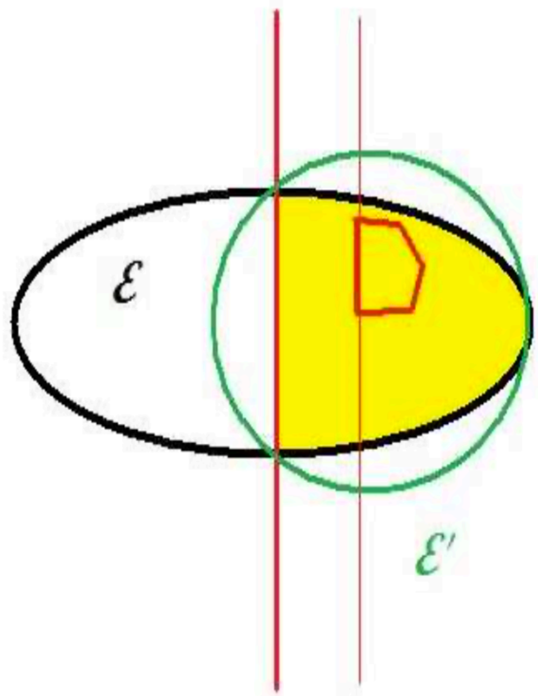
$$z^{k+1} = z^k - \frac{1}{m+1} \frac{P^k a}{\sqrt{a^T P^k a}}$$

$$P^{k+1} = \frac{m^2}{m^2-1}\left(P^k - \frac{2}{m+1}\frac{P^k a a^T P^k}{a^T P^k a}\right)$$

$k = k + 1$

end

At most $O(m^2 \log \sigma)$ iterations.



Notes on Ellipsoid Method

- start from the origin (or arbitrary point),
- the initial ball contains the solution set,
- uses separating hyperplanes (cut generation),
- the new ellipsoid contains the half-ellipsoid containing the solution set,
- finds an optimal solution after finite number of iterations.
- for the original problem, with blown-up solution set a rounding procedure is needed
- polynomial complexity, max $O(m^4 \log \sigma)$, arithmetic operations,
- no efficient implementation exists, all cases like the worst case

The Projective IPM

- Karmarkar reformulated the problem such that the optimal value $c^T x^* = 0$
- Defines the **Barrier Function**:

$$f(x) = n \ln c^T x - \sum_{j=1}^{n} \ln x_j = \sum_{j=1}^{n} \ln\left(\frac{c^T x}{x_j}\right)$$

- **Guarantee:** $f(x^k) - f(x^{k+1}) \geq \delta$
- As $x_j \to 0$, $f \to +\infty$. **Iterates does not move along the "edge", but in the interior (of what?)**
- Complexity of $O(n^{3.5}L^2)$ to reach $c^T x^k / c^T x^0 < 2^{-L}$
- Barrier function defined by Frisch in 1955 logarithmic barrier function (Ragnar Frisch 1895—1973)

$$f(x) = c^T x - \sum_{j=1}^{n} \ln x_j$$

Interior Point Method
（内点法）

北大教材：7.3章

Some Milestones in the History of LO

- 1947 simplex method – Dantzig; still very efficient
- 1972 exponential example – Klee–Minty – theoretical $O(2^n)$
- 1979 ellipsoid method $O(n^6L^2)$ – Khachian; extended to QP
- 1979–1985 ellipsoid method and combinatorial optimization
- 1984 projective IPM $O(n^{3.5}L^2)$ – Karmarkar – efficient in practice!?
- 1989 $O(n^3L)$ for IPMs – Renegar, Gonzaga, Roos, Vial — best complexity
- 1989 Primal–Dual IPMs – Kojima … – dominant since then
- 1996-2000 Volumetric center cutting plane IPMs — by Vaidya, Atkinson and Anstreicher
  IPMs completely replace ellipsoid methods
- 2004 Klee-Minty example for IPMs —> The complexity upper bound is tight

The complexity of IPM

- Textbook: need $O(n \log \frac{\mu_0}{\epsilon})$ iterations to reduce from μ_0 to ϵ
- Best known result $O(\sqrt{n} \log \frac{\mu_0}{\epsilon})$ iterations —For example: *Florian A. Potra, Stephen J. Wright, Interior-point methods, Journal of Computational and Applied Mathematics, Volume 124, Issues 1–2, 2000, Pages 281-302.*
- Many studies to see whether better bounds exist, but the answer is no; the complexity is tight — *A tight iteration-complexity upper bound for the MTY predictor-corrector algorithm via redundant Klee-Minty cubes*

Basic IPM

$$(P) \qquad \min\{c^T x : Ax = b, x \geq 0\},$$

$$(D) \qquad \max\{b^T y : A^T y + s = c, s \geq 0\}$$

$$X := \mathrm{diag}(x_1, x_2, \ldots, x_n), \quad S := \mathrm{diag}(s_1, s_2, \ldots, s_n)$$

$$F(x, y, s) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ XSe \end{bmatrix}$$

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

Naive Newton's Method 求解非线性方程组:

$$F(z)=0 \implies F(z^k)+\nabla F(z^k)(z-z^k)=0$$

$$\implies z^{k+1}=z^k-\nabla F(z^k)^{-1}F(z^k)$$

$$\boxed{z^{k+1} \leftarrow z^k - \alpha^k \nabla F(z^k)^{-1}F(z^k)}$$

$$\begin{cases} z^{k+1} \leftarrow z^k + \alpha^k \Delta z^k \\ \Delta z^k = -\nabla F(z^k)^{-1}F(z^k) \end{cases}$$

Subproblem: solving a system of linear equations

$$\boxed{\nabla F(z^k)\Delta z = -F(z^k)}$$

Basic idea: use Newton's Method to solve

$$F(x, y, s) = 0$$
$$x \geq 0,\ s \geq 0$$

$$\nabla F(x^k, y^k, s^k) \cdot \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} = -F(x^k, y^k, s^k)$$

怎么选α^k?

$$(x^{k+1}, y^{k+1}, s^{k+1}) \leftarrow (x^k, y^k, s^k) + \alpha^k \cdot (\Delta x^k, \Delta y^k, \Delta s^k)$$

如何保证
$x^{k+1} \geq 0,\ s^{k+1} \geq 0$

Basic idea: use Newton's Method to solve

$$F(x, y, s) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ XSe \end{bmatrix} = 0, \quad x \geq 0,\ s \geq 0$$

$$\mathscr{F} := \{(x, y, s) : Ax = b,\ A^T y + s = c,\ x \geq 0,\ s \geq 0\}$$

$$\mathscr{F}^o := \{(x, y, s) \in X : x > 0 \text{ and } s > 0\}$$

Naive Primal-Dual Interior-Point Algorithm

- start with a point $(x, y, s) \in \mathscr{F}^0$
- generate a new point $(x, y, s) + \alpha \cdot (\Delta x, \Delta y, \Delta s)$, where $(\Delta x, \Delta y, \Delta s)$ is computed as a Newton's Step, $\alpha > 0$ is selected so that the new iterate is in $\mathscr{F}$
- Repeat

Basic idea: use Newton's Method to solve

$$F(x,y,s) = \begin{bmatrix} Ax-b \\ A^Ty+s-c \\ XSe \end{bmatrix} = 0, \quad x \geq 0,\ s \geq 0$$

Affine scaling direction:

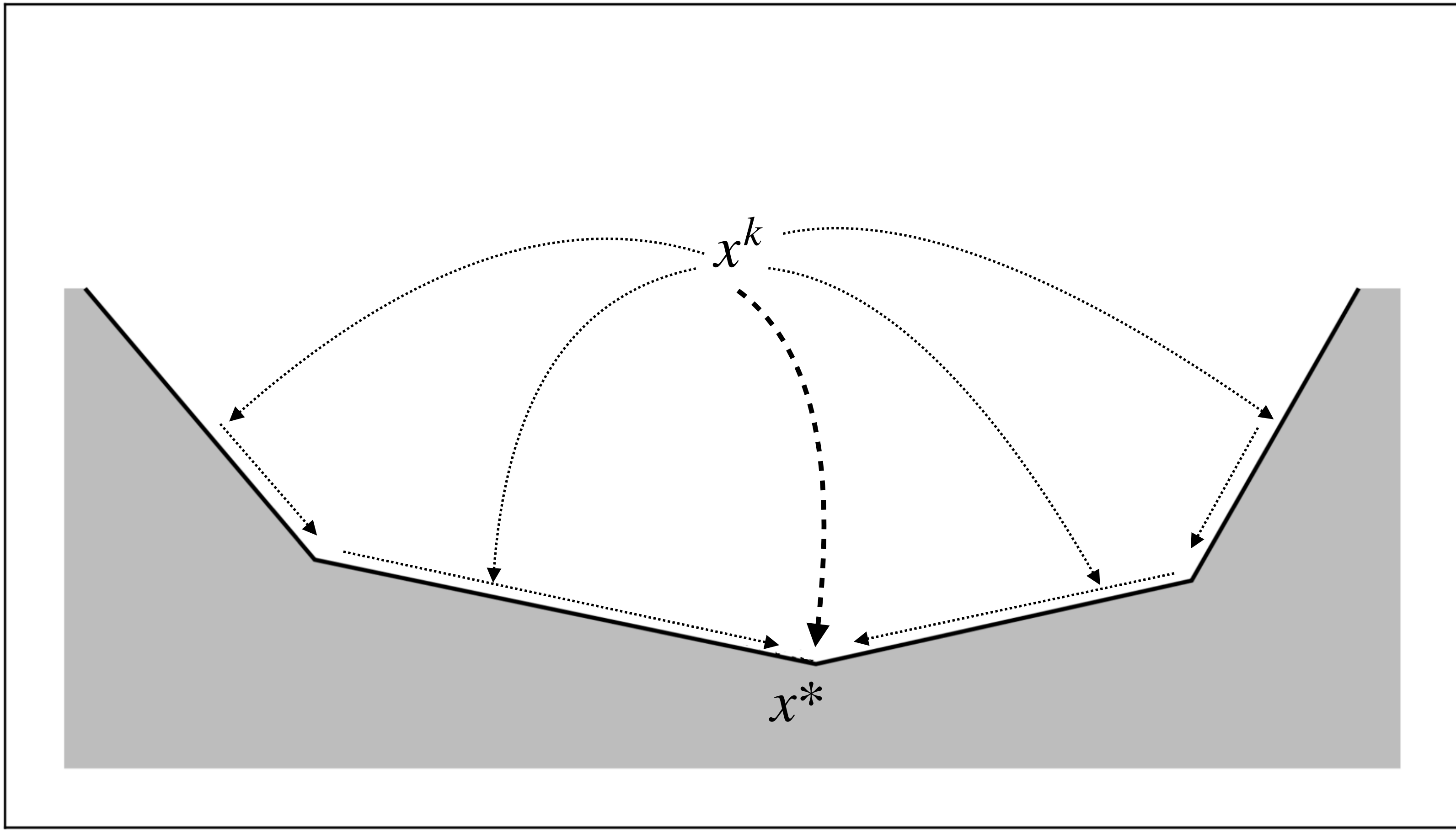
$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L_s & 0 & L_x \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} b-Ax^k \\ -A^Ty^k-s^k+c \\ -X^kS^ke \end{bmatrix} \begin{matrix} \rightarrow \\ \rightarrow \\ \rightarrow \end{matrix} \begin{bmatrix} r_p \\ r_d \\ r_c \end{bmatrix}$$

$$L_s = S^k,\ L_x = X^k$$

$$
\begin{cases}
\Delta y = (AL_s^{-1}L_xA^{\mathrm{T}})^{-1}(r_p + AL_s^{-1}(L_xr_d - r_c)), \\
\Delta s = r_d - A^{\mathrm{T}}\Delta y, \\
\Delta x = -L_s^{-1}(L_x\Delta s - r_c),
\end{cases}
$$

$$
\boxed{(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k, y^k, s^k) + \alpha_k(\Delta x^k, \Delta y^k, \Delta s^k)}
$$

- $AL_s^{-1}L_xA^{\mathrm{T}}$ 是对称矩阵，当 A 满秩时，$AL_s^{-1}L_xA^{\mathrm{T}}$ 正定.
- why we maintain $(x, y, s) \in \mathscr{F}^o$? meaning $XS > 0$
- However, the optimal solution must satisfy $SX = 0$



1. They bias the search direction toward the interior of the nonnegative orthant $(x, s) \geq 0$, so that we can move further along the direction before one of the components of (x, s) becomes negative.

2. They prevent the components of (x, s) from moving "too close" to the boundary of the nonnegative orthant.

Basic idea: add perturbation $\tau > 0$

$$F(x, y, s; \tau) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ XSe - \tau e \end{bmatrix} = 0, \quad x \geq 0,\ s \geq 0$$

$$\begin{aligned} Ax - b &= 0 \\ A^T y + s - c &= 0 \\ x_j s_j &= \tau, j = 1, \ldots, n \\ x \geq 0,\ s &\geq 0 \end{aligned}$$

$\Rightarrow$ Drive $\tau \to 0$ over iteration, so the iterates approaches optimal solution.

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L_s & 0 & L_x \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} b - Ax^k \\ -A^T y^k - s^k + c \\ \tau^k e - X^k S^k e \end{bmatrix} \to \begin{bmatrix} r_p \\ r_d \\ r_c \end{bmatrix}$$

定义 7.6 (中心路径) 给定参数 $\tau > 0$，点 (x_τ, y_τ, s_τ) 满足如下方程：

$$\begin{aligned} Ax &= b, \\ A^{\mathrm{T}}y + s &= c, \\ x_i s_i &= \tau, \quad i = 1, 2, \cdots, n, \\ x > 0, s &> 0. \end{aligned} \tag{7.3.8}$$

则称单参数曲线

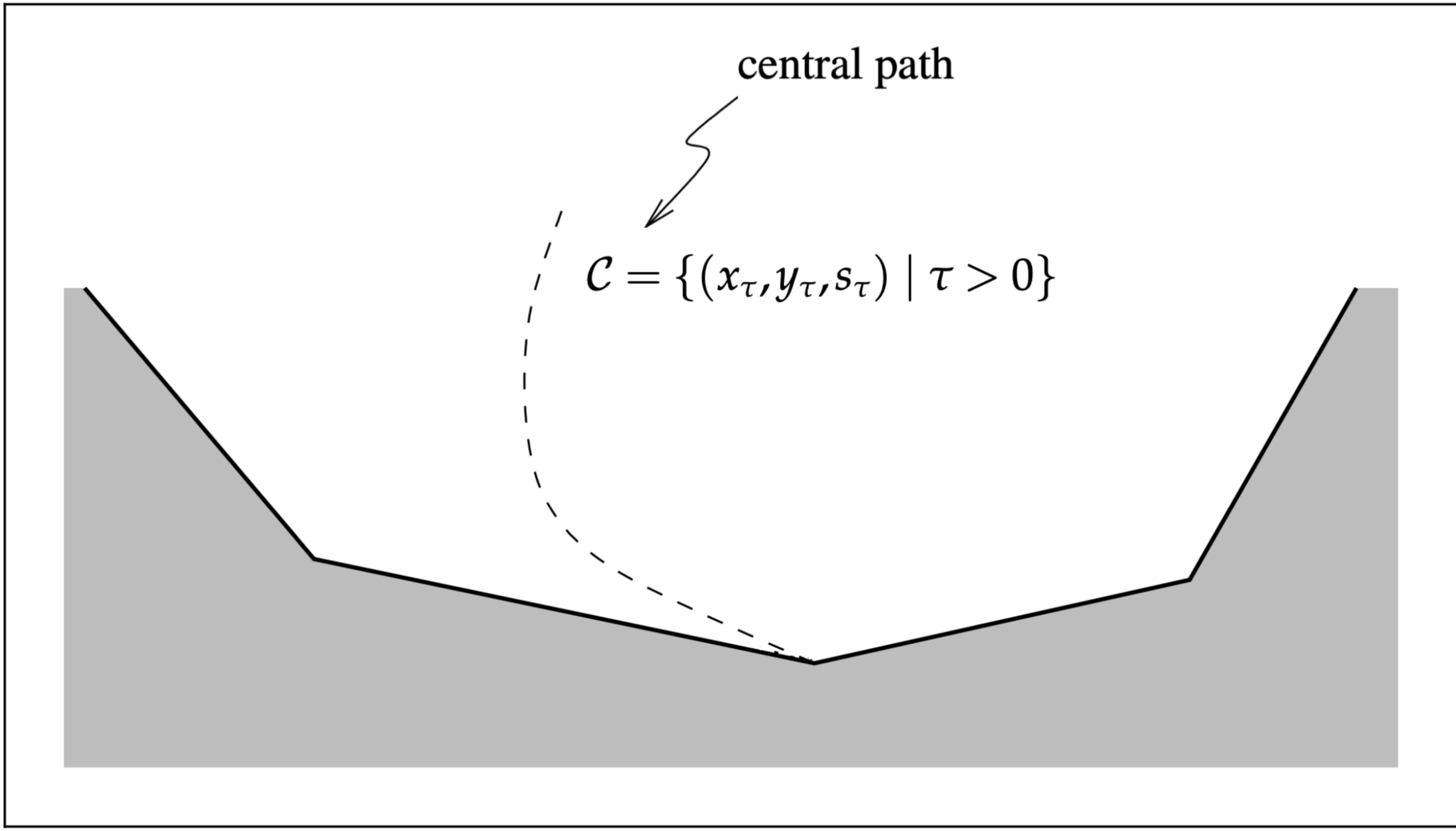
$$\mathcal{C} = \{(x_\tau, y_\tau, s_\tau) \mid \tau > 0\} \tag{7.3.9}$$

为**中心路径**，称方程(7.3.8)为**中心路径方程**.

实际上，从罚函数角度来说，可以证明方程(7.3.8)实际是罚函数形式优化问题

$$\min_x \quad c^{\mathrm{T}}x - \tau \sum_{i=1}^n \ln x_i, \quad \text{s.t.} \quad Ax = b$$

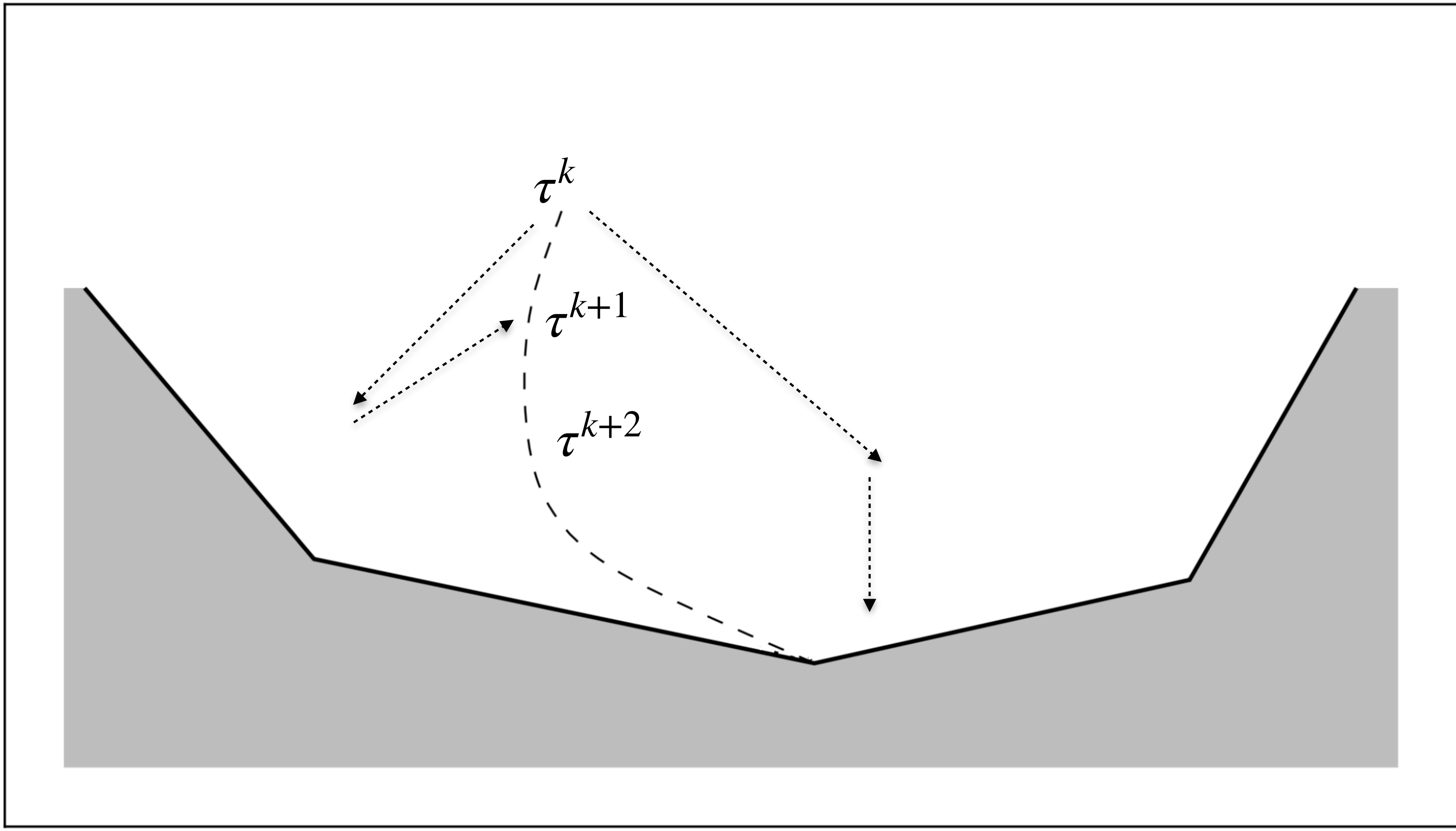
的最优性条件.


central path
$\mathcal{C} = \{(x_\tau, y_\tau, s_\tau) \mid \tau > 0\}$

Basic Primal-Dual Interior-Point Algorithm

- start with a point $(x, y, s) \in \mathcal{F}^o$
- generate a new point $(x, y, s) + \alpha \cdot (\Delta x, \Delta y, \Delta s)$, where $(\Delta x, \Delta y, \Delta s)$ is computed as a Newton's Step to
$$F(x, y, s; \tau^k) = 0,$$
$\alpha > 0$ is selected so that the new iterate is in $\mathcal{F}^o$
- $\tau^{k+1} \leftarrow \sigma\tau^k, \quad \sigma \in (0,1)$
- Repeat

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L_s & 0 & L_x \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} b - Ax^k \\ -A^T y^k - s^k + c \\ \sigma\tau^k e - X^k S^k e \end{bmatrix}$$



- Hard to choose $\{\tau^k\}$

Average distance/average duality gap:

$$\mu(x,s) := \frac{1}{n}\sum_{j=1}^{n} x_j s_j = \frac{x^T s}{n}$$

a simple primal-dual algorithm applies Newton's method to $F(x,y,s;\sigma\mu^k)=0$ where

$$x_j s_j = \sigma\mu(x^k,s^k),\ j=1,\ldots,n;\ \ \sigma\in[0,1]$$

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L_s & 0 & L_x \end{bmatrix}\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} b - Ax^k \\ -A^T y^k - s^k + c \\ \sigma\mu^k e - X^k S^k e \end{bmatrix} \begin{matrix} \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{matrix} \begin{bmatrix} r_p \\ r_d \\ r_c \end{bmatrix}$$

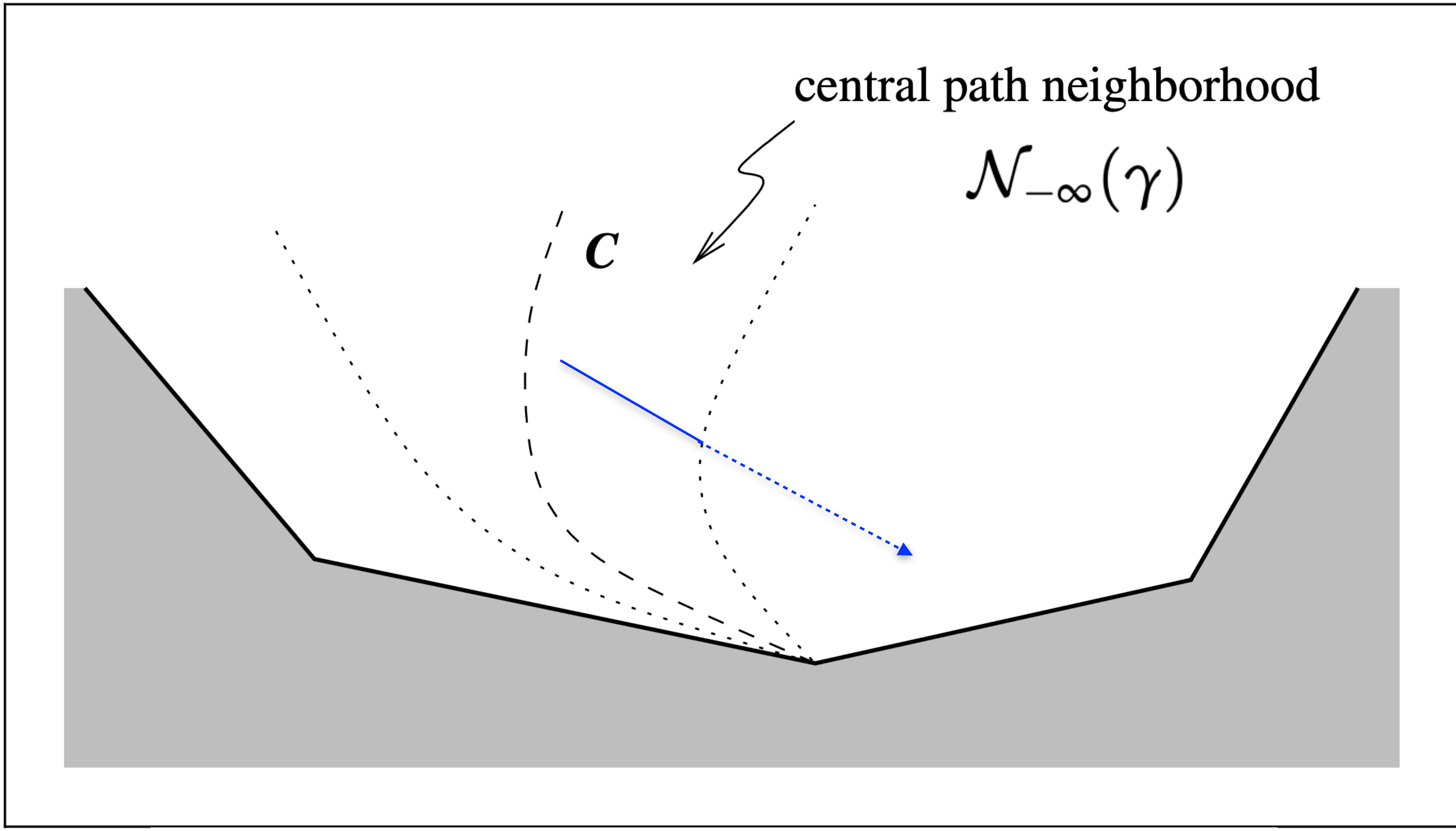
Generic Primal-Dual Interior-Point Algorithm

- start with a point $(x, y, s) \in \mathscr{F}^o$
- generate a new point $(x, y, s) + \alpha \cdot (\Delta x, \Delta y, \Delta s)$, where $(\Delta x, \Delta y, \Delta s)$ is computed as a Newton's Step to $F(x, y, s; \sigma\mu^k) = 0$, $\alpha > 0$ is selected so that the new iterate is in $\mathscr{F}^o$
- Repeat

◆ If $\sigma = 0$, this coincides with the naive primal-dual algorithm

◆ If $\sigma = 1$, this coincides with all $x_j s_j$'s equal to the average duality gap at the current point. Too conservative to aim for an optimal solution, as the average distance may not decrease at all

◆ Best to stay closer to the central path



$$\mathcal{N}_{-\infty}(\gamma) = \{(x,y,s) \in \mathcal{F}^\circ \mid x_i s_i \geqslant \gamma\mu,\ \forall\, i\}.$$

- How to main $\{x^k, y^k, s^k\}$ within the neighborhood?

引理 7.3 设 $(x,y,s) \in \mathcal{N}_{-\infty}(\gamma)$，记

$$(x(\alpha),y(\alpha),s(\alpha)) = (x,y,s) + \alpha(\Delta x,\Delta y,\Delta s).$$

则对任意的 $\alpha \in \left[0, 2^{3/2}\gamma \frac{1-\gamma}{1+\gamma}\frac{\sigma}{n}\right]$，有

$$(x(\alpha),y(\alpha),s(\alpha)) \in \mathcal{N}_{-\infty}(\gamma).$$

算法 7.7 路径追踪算法

---

1. 选取初值 $(x^0,y^0,s^0) \in \mathcal{F}^\circ$，参数 $0 < \gamma < \sigma < 1$，$k \leftarrow 0$.
2. **while** 未达到收敛准则 **do**
3. 　求解方程(7.3.5)得到更新 $(\Delta x,\Delta y,\Delta s)$.
4. 　选取最大的 $\alpha \in (0,1]$ 使得下一步迭代点落在 $\mathcal{N}_{-\infty}(\gamma)$ 内，记为 α_k.
5. 　更新 $(x^{k+1},y^{k+1},s^{k+1}) = (x^k,y^k,s^k) + \alpha_k(\Delta x,\Delta y,\Delta s)$.
6. 　$k \leftarrow k+1$.
7. **end while**

Complexity

Binary Input Length:

$$L = mn + n + \sum_{i=1}^{n}\sum_{j=1}^{m} \lceil \log_2(|a_{ij}| + 1) \rceil + \sum_{i=1}^{n} \lceil \log_2(|c_j| + 1) \rceil$$

Termination Criterion:

$$(x^k)^T s^k < 2^{-L}$$

$$\begin{aligned} \Longrightarrow \quad c^T x^k - c^T x^* &= c^T x^k - b^T y^* \\ &\le c^T x^k - b^T y^k \\ &= c^T x^k - (Ax^k)^T y \\ &= (c - A^T y^k)^T x^k = (s^k)^T x^k \\ &\le 2^{-L} \end{aligned}$$

Tolerance: in binary representation 2^{-L} has the number of digits bounded in L (namely, $L+$constant)

Complexity

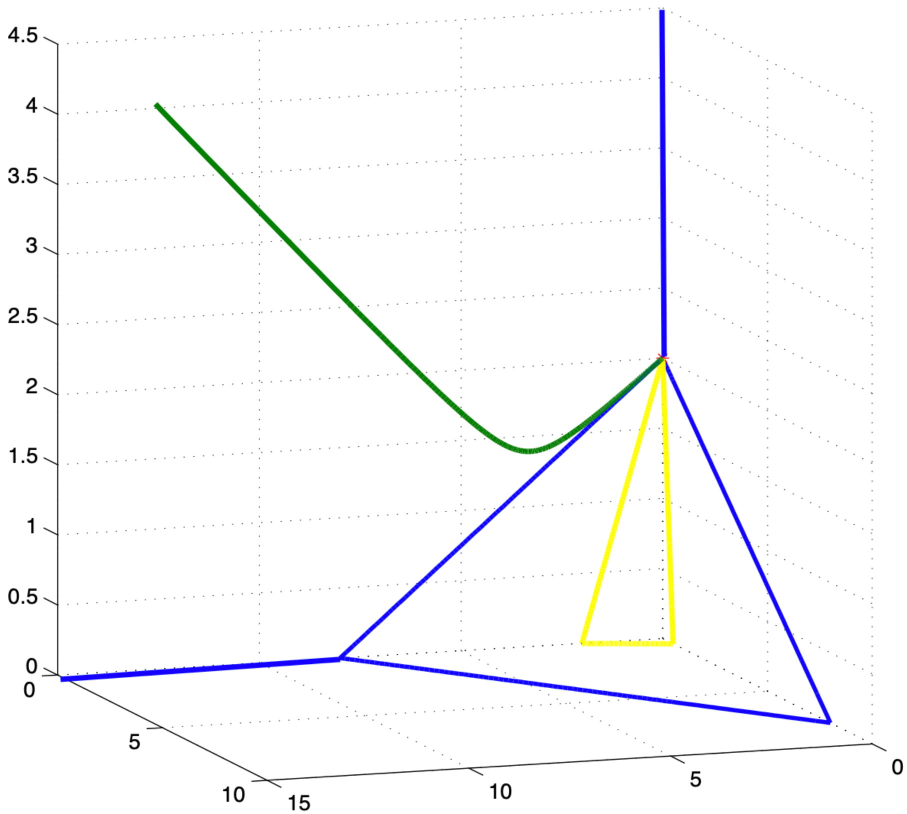
定理 7.10 (原始 – 对偶算法的收敛性) 给定参数 $0<\gamma<\sigma<1$，设 $\mu_k=\dfrac{(x^k)^{\mathrm{T}}s^k}{n}$ 为算法7.7 产生的**对偶间隙**，且初值 $(x^0,y^0,s^0)\in\mathcal{N}_{-\infty}(\gamma)$，则存在与维数 n 无关的常数 c，使得对任意 k 有

$$\mu_{k+1}\leqslant\left(1-\frac{c}{n}\right)\mu_k.$$

更进一步地，对任意给定的精度 $\varepsilon\in(0,1)$，存在迭代步数 $K=\mathcal{O}\left(n\ln\dfrac{1}{\varepsilon}\right)$ 使得

$$\mu_k\leqslant\varepsilon\mu_0,\quad\forall k\geqslant K.$$

- Let $\varepsilon=n2^{-L}$, the number of iterations needed $O(nL)$
- With line search (changing α), best known: $O(\sqrt{n}L)$
- At each iteration, solving linear equations.



- IPMs generate a sequence of strictly positive (interior w.r.t) the positive orthant.
- The iterates are strictly feasible primal and dual solutions.
- The iterates are in the <u>relative interior</u> of the feasible sets.
- Complementarity is reached only at optimality.
- The duality gap decreases monotonically.

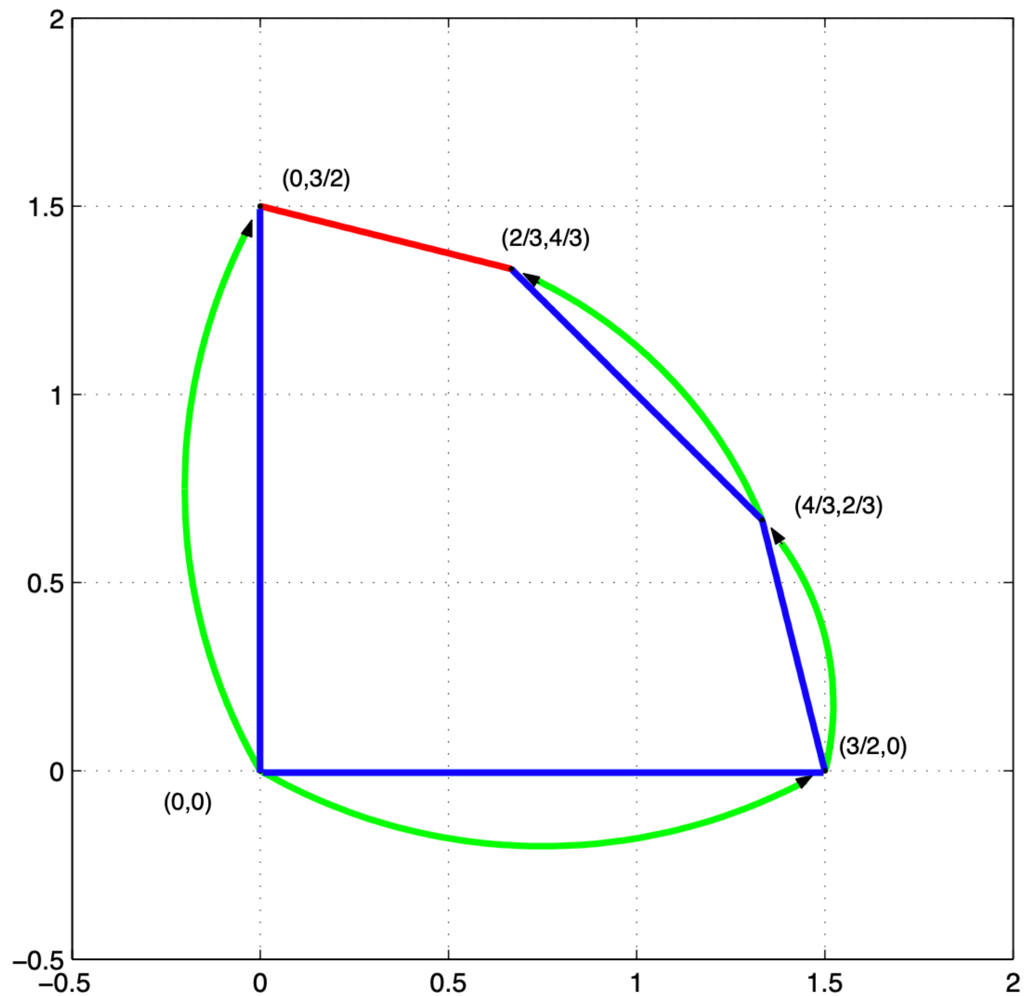
- Simplex methods generate an optimal basic solution.

- In case of degeneracy one does not know which optimal basic solution will be found.

- In case of degeneracy the found optimal basic solution is not strictly complementary.

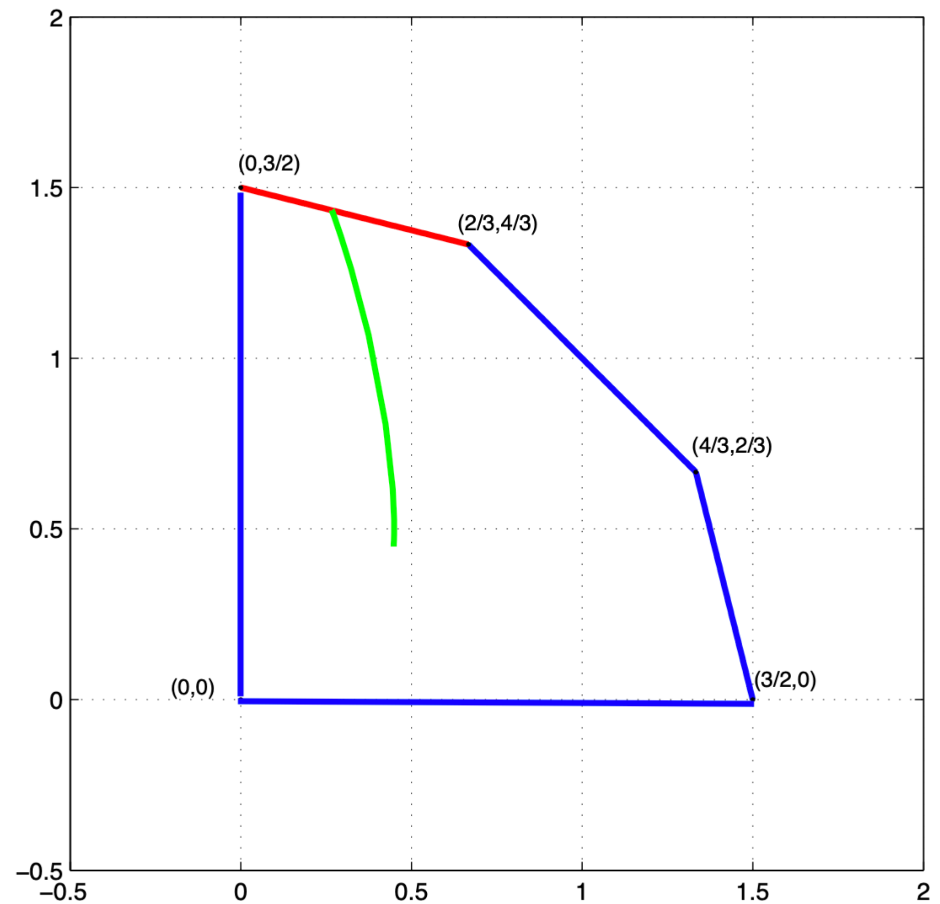
degenerate solution
$\Longrightarrow$ not strictly complementary

- To find a strictly complementary solution from an optimal basis is not easy, it is just as difficult as solving the original problem.



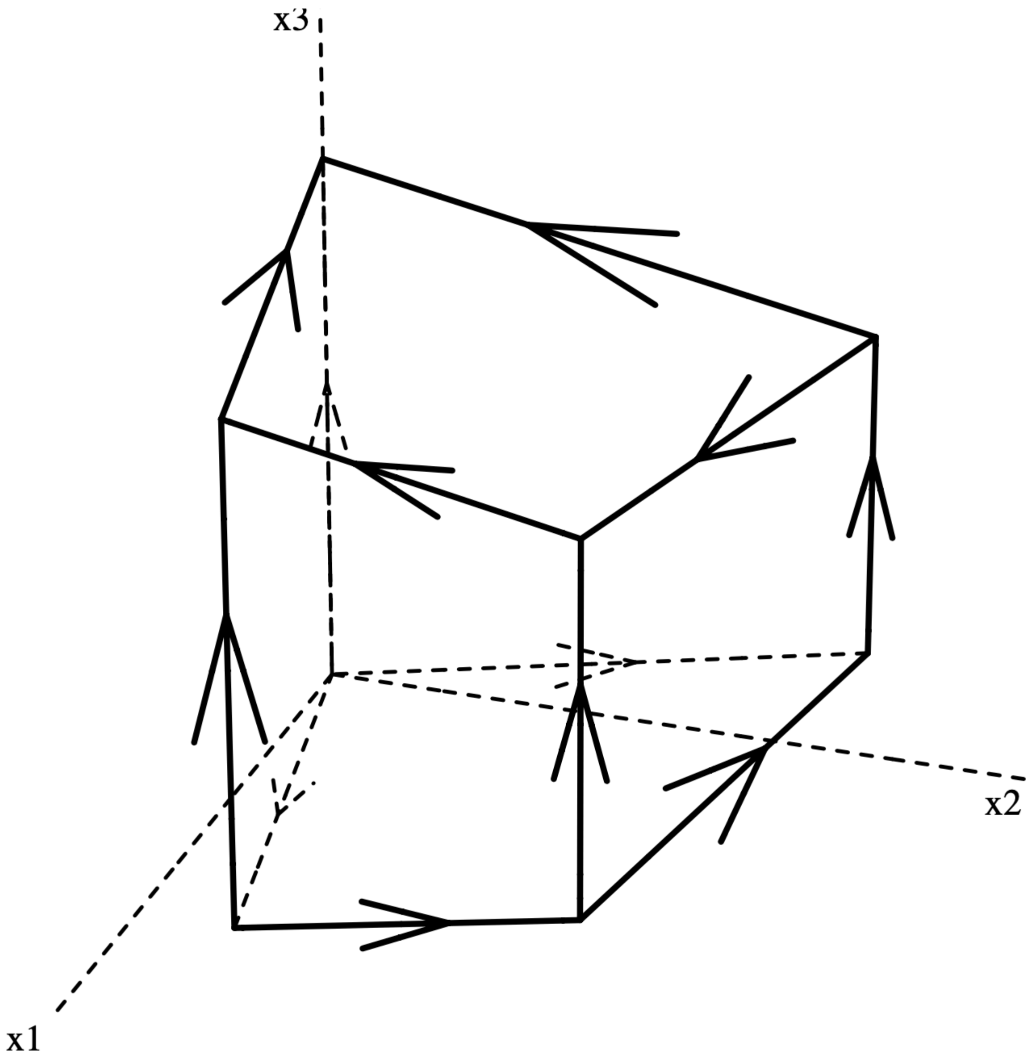
Primal Simplex path and solution(s)

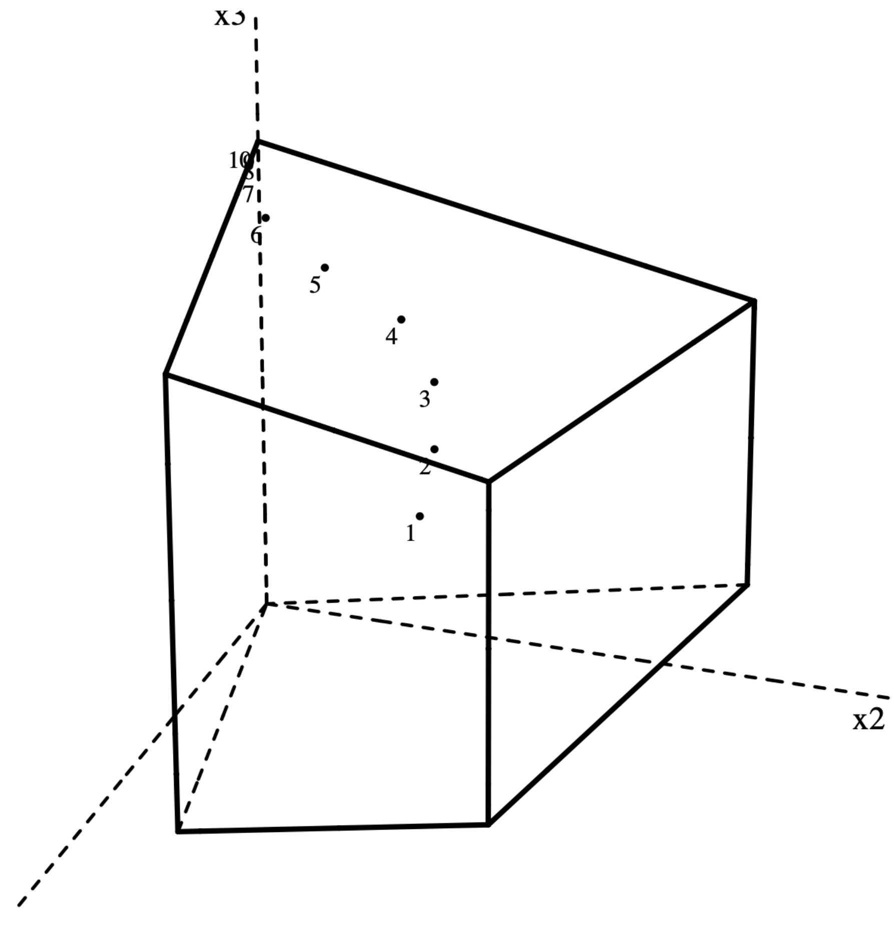
- IPMs produce an ε optimal solution, i.e. a feasible solution pair for which $c^T x - b^T y < \epsilon$
- In case of degeneracy IPMs converge to the analytic centre of the optimal face. (not a pain for IPM)
- Starting from a sufficiently precise solution an exact strictly complementary solution can readily be identified.
- An optimal basis can be found by using at most n additional pivots.

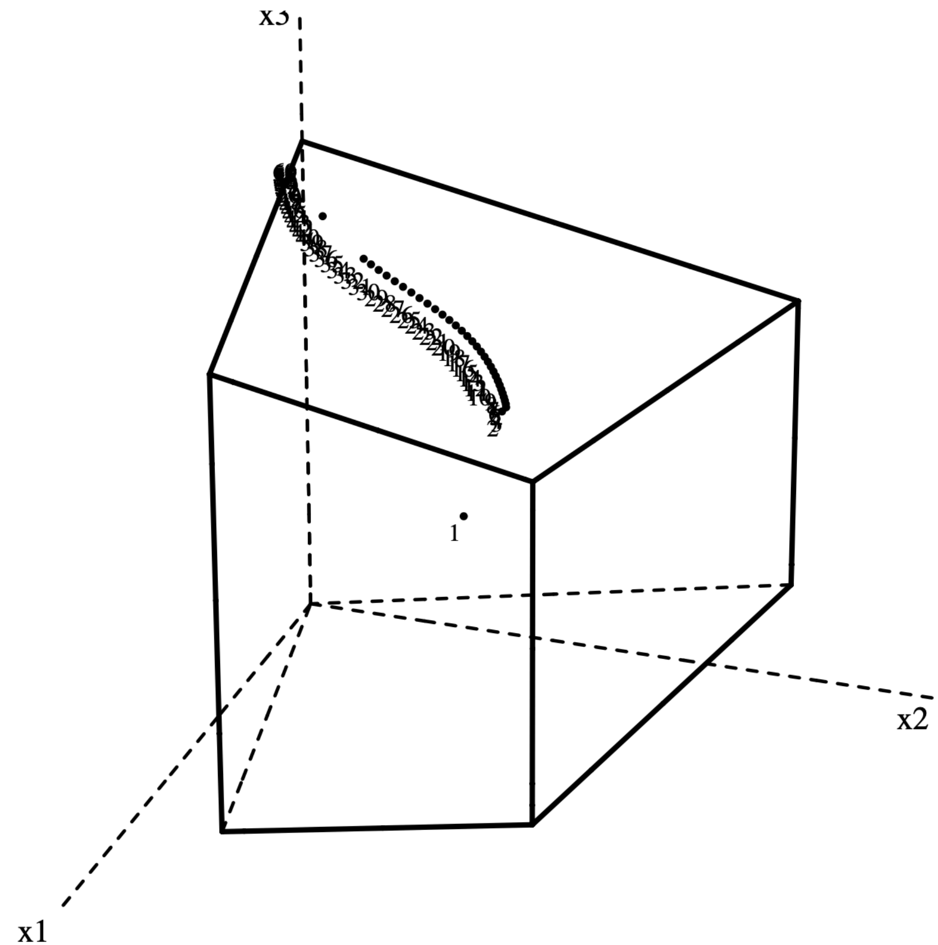


IPM Path – primal

3D Klee-Minty's Example




x3
10
8
7
6
5
4
3
2
1
x2


x3
1
2
x2
x1