Interior Point Method

Junzhe He

School of Computer Science and Technology
Zhejiang University

October 29, 2025

Contents

History

- In 1967, Ilya I. Dikin proposed *Affine Scaling* method for linear programming.
- In 1984, Karmakar at AT&T "invented" interior point method.
- In 1986, Affine scaling was "invented" by R.J. Vanderbei *et. al.* at IBM + AT&T , which far outperformed simplex method and Karmakar's method in practice.
- In 1988, R. Monteiro and I. Adler at UCB proposed *primal-dual interior point* methods.
- In 1992, V.K. Mehrotra proposed a *predictor-corrector* variant of primal-dual interior point method, which is the most widely used method in practice today.
- In 1994, Nesterov and Nemirovski developed the theory of *self-concordant* barrier functions.

Intuition

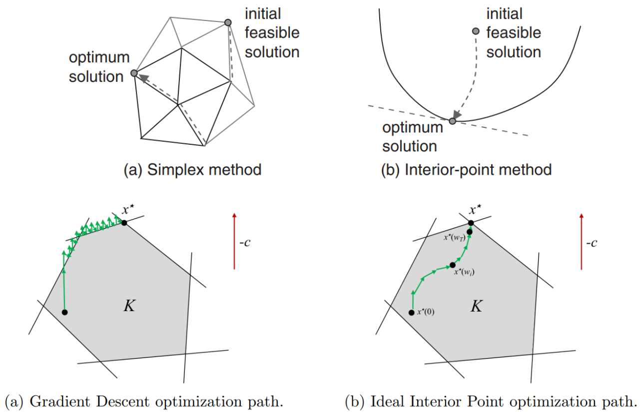


Figure 1: Intuition of Interior Point Method.[1]

[1] Image from Wang, L. T., Chang, Y. W., & Cheng, K. T. T. . (2009). *Electronic design automation: synthesis, verification, and test*. Morgan Kaufmann. Chapter 4.5.4.1. and Musco C. (2018). *Lecture 17: Interior Point Methods*. Retrieved from `https://www.cs.princeton.edu/courses/archive/fall18/cos521/Lectures/lec17.pdf`.

Three Methods

- Affine-Scaling
- Potential function reduction
- Path-Following

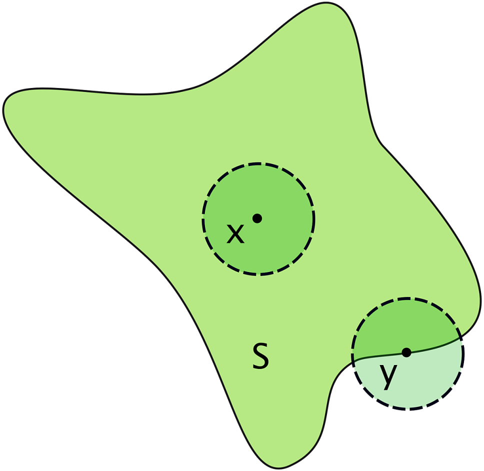


Figure 2: Illustration of interior of a set.[2]

[2] Image from https://en.wikipedia.org/wiki/Interior_(topology)

What is the interior of a LP problem?

(P)

$$\begin{aligned} \min\ & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{aligned}$$

(D)

$$\begin{aligned} \max\ & \boldsymbol{b}^\top \boldsymbol{y} \\ \text{s.t. } & \boldsymbol{A}^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ & \boldsymbol{s} \geq 0 \end{aligned}$$

Affine-Scaling

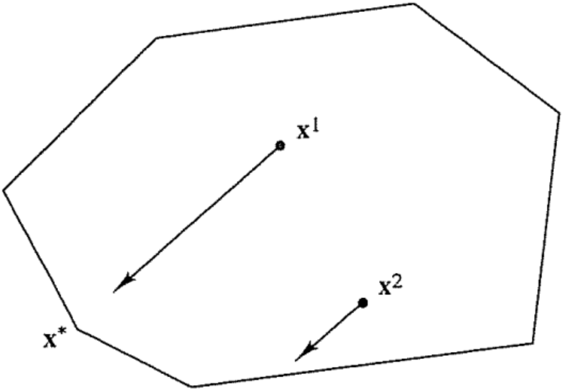


Figure 3: Since x^1 is near the center of the polytope, we can improve the solution substantially by moving it in a direction of steepest descent. But if an off-center point x^2 is so moved, it will soon be out of the feasible domain.[3]

[3] Image from Fang, S. C., & Puthenpura, S. (1993). *Linear optimization and extensions: theory and algorithms*. Prentice-Hall, Inc.. 6.1

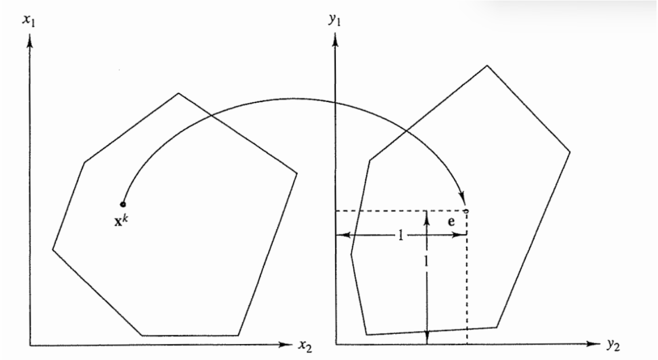


Figure 4: Illustration of *Affine Scaling transformation*.[4]

[4] Image from Fang, S. C., & Puthenpura, S. (1993). *Linear optimization and extensions: theory and algorithms*. Prentice-Hall, Inc.. 7.1.1.

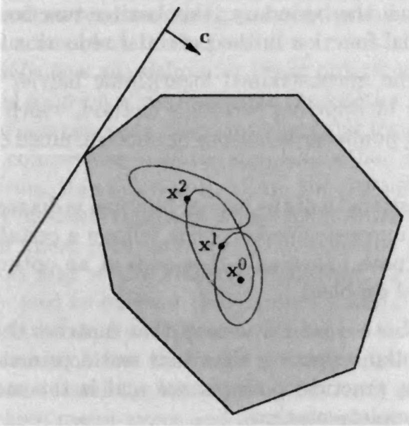


Figure 5: Illustration of Affine Scaling method iteration.[5]

[5] Image from Bertsimas, D., & Tsitsiklis, J. N. (1997). *Introduction to linear optimization*. Belmont, MA: Athena Scientific. 9.1

Input: A feasible solution $\boldsymbol{x}^0$ to (P)
for $k = 0, 1, 2, \ldots$ **do**
Given the current solution $\boldsymbol{x}^k$, let

$$\boldsymbol{X}^k = \operatorname{diag}(x_1^k, x_2^k, \ldots, x_n^k)$$
$$\boldsymbol{p}^k = -P_{A\boldsymbol{X}^k}(\boldsymbol{X}^k c)$$
$$\boldsymbol{d}^k = \frac{\boldsymbol{X}^k \boldsymbol{p}^k}{\|\boldsymbol{p}^k\|}$$

if $\|\boldsymbol{p}^k\| < \epsilon$ **then**
Stop and **return** $\boldsymbol{x}^k$ as an approximate optimal solution.
end
Update the solution: $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \alpha^k \boldsymbol{d}^k$.
end

$$(P(\mu)) \qquad \begin{aligned} &\min F_\mu(x) := \boldsymbol{c}^\top \boldsymbol{x} + \mu B(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

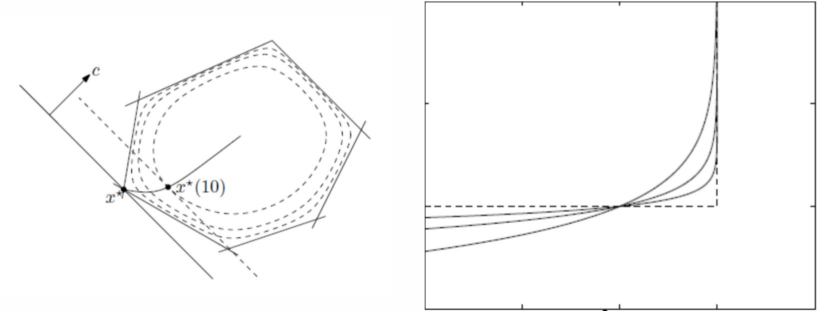


Figure 6: Barrier Function Illustration.[6]

[6] Image from https://www.stat.cmu.edu/~ryantibs/convexopt-S15/scribes/15-barr-method-scribed.pdf.

Why we choose the logarithmic barrier function?

$$B(x) := -\sum_{i=1}^{n} \ln x_i$$

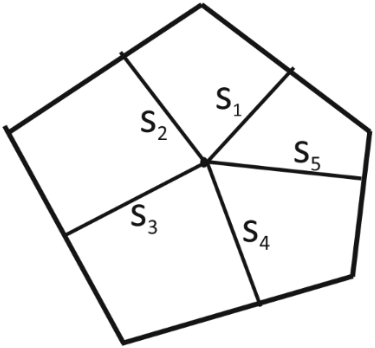


Figure 7: The Analytic Center.[7]

[7] Image from Luenberger, D. G., & Ye, Y. (2021). *Linear and nonlinear programming*. Reading, MA: Addison-wesley. 5.4

Why we choose the logarithmic barrier function?

$$B(x) := -\sum_{i=1}^{n} \ln x_i$$

"$\ln$" function is *Self-concordant*. A function $f : \mathbb{R} \to \mathbb{R}$ is self-concordant when

- f is convex and three times differentiable.
- $|f'''(x)| \leq 2(f''(x))^{3/2}$ for all x in the domain of f.

Newton’s Method for Self-concordant Functions

Theorem [Nesterov & Nemirovski, 1994][7]

Damped Newton’s Method (with line search step length) could find an ϵ-approximate solution to the minimization of a self-concordant function f in

$$C(f(\boldsymbol{x}^0) - f(\boldsymbol{x}^*)) + 2\log\log\frac{1}{\epsilon}$$

steps.

[7] Nesterov, Y., & Nemirovskii, A. (1994). *Interior-point polynomial algorithms in convex programming*. SIAM.

Barrier Problem

$$(P(\mu)) \qquad \begin{aligned} \min\ & F_\mu(x) := \boldsymbol{c}^\top \boldsymbol{x} - \mu \sum_{i=1}^{n} \ln x_i \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

ZHEJIANG UNIVERSITY

Barrier Problem

$$(P(\mu)) \qquad \min F_\mu(x) := \boldsymbol{c}^\top \boldsymbol{x} - \mu \sum_{i=1}^{n} \ln x_i$$
$$\text{s.t. } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

KKT conditions

$$\begin{cases} \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ \boldsymbol{A}^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ x_i s_i = \mu, \quad i = 1, 2, \ldots, n \end{cases}$$

Central paths

- Primal central path: $\{\boldsymbol{x}(\mu) : \mu > 0\}$;
- Primal-Dual central path: $\{\boldsymbol{x}(\mu), \boldsymbol{y}(\mu), \boldsymbol{s}(\mu) : \mu > 0\}$.

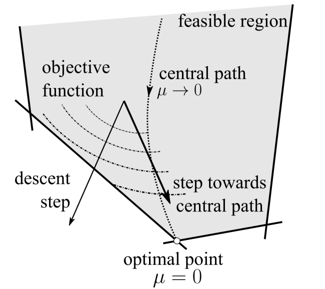


Figure 7: The central path defines a set of well-centered points located far from the feasible region boundaries, which leads to optimal solution when $\mu \to 0$.[8]

[8] Image from Bleyer, J. (2018). *Advances in the simulation of viscoplastic fluid flows using interior-point methods*. Computer Methods in Applied Mechanics and Engineering, 330, 368-394..

Primal-Dual Newton equation system

$$\begin{cases} A\Delta x = 0 \\ A^\top \Delta y + \Delta s = 0 \\ S\Delta x + X\Delta s = \mu \mathbf{1}_n - XS\mathbf{1}_n \end{cases}$$

Primal-Dual Newton equation system

$$\begin{cases} A\Delta x = 0 \\ A^\top \Delta y + \Delta s = 0 \\ S\Delta x + X\Delta s = \mu \mathbf{1}_n - XS\mathbf{1}_n \end{cases}$$

Primal-Dual Newton step

$$\begin{cases} \Delta y = -(AS^{-1}XA^\top)^{-1}AS^{-1}(\mu \mathbf{1}_n - XS\mathbf{1}_n) \\ \Delta s = -A^\top \Delta y = (AS^{-1}XA^\top)^{-1}AS^{-1}(\mu \mathbf{1}_n - XS\mathbf{1}_n) \\ \Delta x = -x + \mu S^{-1}\mathbf{1}_n - S^{-1}X\Delta s \end{cases}$$

We define $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ to be a β-*approximate solution* of $P(\mu)$ if it satisfies

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{A}^\top \boldsymbol{y} + \boldsymbol{s} &= \boldsymbol{c} \\ \left\| \frac{1}{\mu} XS\mathbf{1}_n - \mathbf{1}_n \right\| &\leq \beta, \quad i = 1, 2, \ldots, n \end{aligned}$$

Theorem (Explicit Quadratic Convergence of Newton's Method)

Suppose that $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ is a β-approximate solution of $P(\mu)$ with $\beta < \frac{1}{2}$. Let $(\boldsymbol{x} + \Delta \boldsymbol{x}, \boldsymbol{y} + \Delta \boldsymbol{y}, \boldsymbol{s} + \Delta \boldsymbol{s})$ be the updated solution after one primal-dual Newton step. Then, $(\boldsymbol{x} + \Delta \boldsymbol{x}, \boldsymbol{y} + \Delta \boldsymbol{y}, \boldsymbol{s} + \Delta \boldsymbol{s})$ is a

$$\left(\frac{1 + \beta}{(1 - \beta)^2} \right) \beta^2$$

-approximate solution of $P(\mu)$.

Input: A $\frac{1}{3}$-approximate feasible solution $(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{s}^0)$, parameter μ^0.
for $k = 0, 1, 2, \ldots$ **do**

    if $\max\{\|\boldsymbol{A}\boldsymbol{x}^k - \boldsymbol{b}\|, \|\boldsymbol{A}^\top \boldsymbol{y}^k + \boldsymbol{s}^k - \boldsymbol{c}\|, {\boldsymbol{s}^k}^\top \boldsymbol{x}^k\} < \epsilon$ **then**

        Stop and **return** $\boldsymbol{x}^k$ as an approximate optimal solution.

    end

    Shrink:

$$\mu^{k+1} = \alpha \mu^k \text{ for some } \alpha \in (0, 1)$$

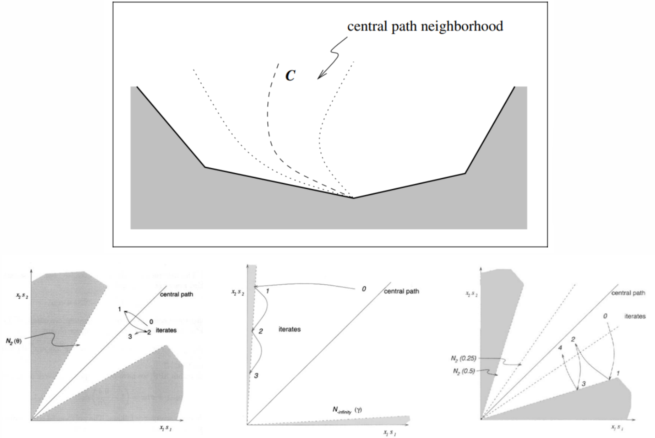
    Compute the primal-dual Newton step $(\Delta \boldsymbol{x}^k, \Delta \boldsymbol{y}^k, \Delta \boldsymbol{s}^k)$ w.r.t. $(\boldsymbol{x}^k, \boldsymbol{y}^k, \boldsymbol{s}^k)$ for μ^{k+1} .

    Update the solution:

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

end

Path Following Methods



Short step path-following **Long step path-following** **Predict-Correct**

Figure 8: Illustration of Path Following Neighborhood.[9]

---

[9] Image fromWright, S. J. (2006). *Numerical optimization*. Springer Science+Business Media, LLC. 14.1. and Wright, S. J. (1997). *Primal-dual interior-point methods*. SIAM. Chapter 5..

Williamson, David P. (2014)
ORIE 6300: Mathematical Programming I
Cornell University, Fall 2014,
https://people.orie.cornell.edu/dpw/orie6300/.

Freund, Robert M. (2014)
Primal-Dual Interior-Point Methods for Linear Programming based on Newton's Method
Lecture Notes , http://s3.amazonaws.com/mitsloan-php/wp-faculty/sites/30/2016/12/15031751/Interior-Point-Methods-for-Linear-Optimization.pdf

End

Thanks for your listening.