Interior Point Methods for Linear Programming Problems

Lecture 2

Debdas Ghosh

Contents

- Interior point methods
- Logarithmic barrier method
- Primal-dual interior point method

Interior point methods

- Logarithmic barrier method
- Path-following
- Center method
- Affine scaling
- Potential reduction

Contents

Logarithmic barrier method

Given LPP

$$(\mathrm{P}) \quad \begin{cases} \min & c^T x \\ \text{subject to} & Ax = b, x \geq 0. \end{cases}$$

The log-barrier formulation of (P) is

$$(\mathrm{LBP}) \quad \begin{cases} \min & \phi_B(x;\mu) \equiv \frac{c^T x}{\mu} - \sum_{i=1}^{n} \ln(x_i) \\ \text{subject to} & Ax = b. \end{cases}$$

The first order necessary and sufficient optimality conditions for (LBP) is

$$(\mathrm{LBP\ OC}) \quad \begin{cases} \nabla \phi_B(x;\mu) + \sum_{i=1}^{m} \lambda_i \nabla h_j(x) = 0 \\ Ax = b \end{cases}$$

Logarithmic barrier method

Given LPP

$$\text{(P)} \quad \begin{cases} \min & c^T x \\ \text{subject to} & Ax = b, x \geq 0. \end{cases}$$

The log-barrier formulation of (P) is

$$\text{(LBP)} \quad \begin{cases} \min & \phi_B(x;\mu) \equiv \frac{c^T x}{\mu} - \sum_{i=1}^{n} \ln(x_i) \\ \text{subject to} & Ax = b. \end{cases}$$

The first order necessary and sufficient optimality conditions for (LBP) is

$$\text{(LBP OC)} \quad \begin{cases} \frac{c}{\mu} - X^{-1}e + A^T\lambda = 0 \\ Ax = b. \end{cases}$$

Logarithmic barrier method

Notations

- $\overset{\circ}{\mathcal{F}}(P) = \{x \in \mathbb{R}^n : Ax = b, x > 0\}$
- $\overset{\circ}{\mathcal{F}}(D) = \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c, s > 0\}$

Existence of a minimizer of $\phi_B(x; \mu)$

A necessary and sufficient condition for the existence of a minimizer $\phi_B(x; \mu)$ in $\overset{\circ}{\mathcal{F}}(P)$ is '$\overset{\circ}{\mathcal{F}}(P)$ and $\overset{\circ}{\mathcal{F}}(D)$ nonempty'.

For a given $\mu > 0$, let $(x(\mu), y(\mu), s(\mu))$ be the solution to (PD OCS)

Logarithmic barrier method

Central path

- $\{x(\mu) : \mu > 0\}$ is called the primal central path.
- $\{(x(\mu), y(\mu), s(\mu) : \mu > 0\}$ is called primal-dual central path.

Analytic center

If the feasible region of (P) is bounded, then the central path starts (i.e., when $\mu \to \infty$) from the unique point

$$\text{argmin} \sum_{i=1}^{n} -\ln(x_i),$$

which is called the analytic center of the feasible region.

Central path

If μ decreases, then $c^T x(\mu)$ decreases and $b^T y(\mu)$ increases.

On the central path, $c^T x(\mu) - b^T y(\mu) = n\mu$.

Example 1

Find the central path and analytic center of the LPP:

$$\max 2x_1 + x_2 \text{ subject to } -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1.$$

Examples of central paths

Example 2

Find the central path and analytic center of the LPP:

$$\max 2x_1 + x_2 \;\; \text{subject to} \;\; x_1 \leq 1, x_2 \leq 1.$$

Examples of central paths

Example 2

Find the central path and analytic center of the LPP:

$$\max 2x_1 + x_2 \ \text{ subject to } \ x_1 \leq 1, x_2 \leq 1.$$

Example 3

Find the central path and analytic center of the LPP:

$$\max -x_1 - 3x_2 - 4x_3$$

$$\text{subject to } \ x_1 + x_2 + x_3 = 1$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

Logarithmic barrier method

Questions arise to find the following:

- The method to approximately solve the barrier problem.
- Criteria to terminate the approximate minimization.
- Updating scheme for μ.

Logarithmic barrier method

Log-barrier formulation of the dual of (P) is

$$\min \quad \psi_B(y;\mu) \equiv -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$$

$$\text{subject to} \quad A^T y + s = c.$$

Logarithmic barrier method

Log-barrier formulation of the dual of (P) is

$$\min \quad \psi_B(y;\mu) \equiv -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$$

subject to $\quad A^T y + s = c.$

- $\psi_B(y;\mu) = -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(c_i - a_i^T y)$

Logarithmic barrier method

Log-barrier formulation of the dual of (P) is

$$\min \quad \psi_B(y;\mu) \equiv -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$$

subject to $\quad A^T y + s = c.$

- $\psi_B(y;\mu) = -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(c_i - a_i^T y)$
- $g(y;\mu) \coloneqq \nabla \psi_B(y;\mu) = -\frac{b}{\mu} + AS^{-1}e$

Logarithmic barrier method

Log-barrier formulation of the dual of (P) is

$$\min \quad \psi_B(y;\mu) \equiv -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$$

subject to $\quad A^T y + s = c.$

- $\psi_B(y;\mu) = -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(c_i - a_i^T y)$
- $g(y;\mu) \coloneqq \nabla \psi_B(y;\mu) = -\frac{b}{\mu} + AS^{-1}e$
- $H(y;\mu) \coloneqq \nabla^2 \psi_B(y;\mu) = \sum_{i=1}^{n} \frac{a_i a_i^T}{(c_i - a_i^T y)^2} = AS^{-2}A^T.$

Logarithmic barrier method

Log-barrier formulation of the dual of (P) is

$$\min \quad \psi_B(y;\mu) \equiv -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$$

subject to $\quad A^T y + s = c.$

- $\psi_B(y;\mu) = -\frac{b^T y}{\mu} - \sum_{i=1}^n \ln(c_i - a_i^T y)$
- $g(y;\mu) := \nabla \psi_B(y;\mu) = -\frac{b}{\mu} + AS^{-1}e$
- $H(y;\mu) := \nabla^2 \psi_B(y;\mu) = \sum_{i=1}^n \frac{a_i a_i^T}{(c_i - a_i^T y)^2} = AS^{-2}A^T.$
- Newton direction:

$$d(y;\mu) := -H^{-1} g = (AS^{-2}A^T)^{-1} \left(\frac{b}{\mu} - AS^{-1} e \right)$$

Algorithm of logarithmic barrier method

To solve

$\min -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$ subject to $A^T y + s = c$.

Algorithm of logarithmic barrier method

To solve

$\min -\frac{b^T y}{\mu} - \sum_{i=1}^{n} \ln(s_i)$ subject to $A^T y + s = c$.

Input

- $\varepsilon' > 0 \rightarrow$ accuracy parameter for the duality gap (for the outer loop)
- $\varepsilon'' > 0 \rightarrow$ accuracy parameter for the optimal solution (in the outer loop)
- $\theta \in (0, 1) \rightarrow$ reduction factor of the barrier parameter μ
- $\mu_0 \rightarrow$ initial value of the barrier parameter
- $y^0 \rightarrow$ initial guess to start Newton's iteration

Algorithm of logarithmic barrier method

Algorithm 1

1. Initialize $\mu \leftarrow \mu_0$ and $y \leftarrow y^0$
2. (Outer loop) While $n\mu > \varepsilon'$

   $\mu \leftarrow \theta\mu$

   2.1 (Inner loop) While $\|g(y;\mu)\| > \varepsilon''$

   Compute $d(y;\mu) = -H^{-1}g = (AS^{-2}A^T)^{-1}\left(\frac{b}{\mu} - AS^{-1}e\right)$

   Compute $\alpha = \text{argmin}\left\{\psi_B(y+\alpha d;\mu) : y+\alpha d \in \mathring{\mathcal{F}}(D)\right\}$

   Update $y \leftarrow y + \alpha d$
3. Output $y(\mu)$ is an ε'-optimal point.

An example for the execution of Algorithm 1

Example

Consider to solve the following (dual) problem by logarithmic-barrier method with pure Newton method in the inner loop:

$$\max \quad y_1 + 2y_2$$
$$\text{subject to} \quad y_1 \leq 0, y_2 \leq 0.$$

Find the minimum of this problem taking $\mu_0 = 1$, $\theta = \frac{1}{2}$, $\varepsilon' = 10^{-2}$, $\varepsilon'' = 10^{-1}$ and $y^0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

Complexity analysis

Algorithm 2

1. Initialize $\mu \leftarrow \mu_0$ and $y \leftarrow y^0$
2. (Outer loop) While $n\mu > \varepsilon$

   $\mu \leftarrow (1-\theta)\mu$

   Compute $d(y;\mu) = -H^{-1}g = (AS^{-2}A^T)^{-1}\left(\frac{b}{\mu} - AS^{-1}e\right)$

   2.1 (Inner loop) While $\|d(y;\mu)\|_H > \frac{1}{2}$

   Compute $d(y;\mu) = -H^{-1}g = (AS^{-2}A^T)^{-1}\left(\frac{b}{\mu} - AS^{-1}e\right)$

   Compute $\alpha = \text{argmin}\left\{\psi_B(y+\alpha d;\mu) : y + \alpha d \in \mathring{\mathcal{F}}(D)\right\}$

   Update $y \leftarrow y + \alpha d$
3. Output $y(\mu)$ is an ε-optimal point.

Complexity analysis

Maximum outer iterations

Let z^* be the optimum value of the dual problem of (P). Then, after at most $\frac{1}{\theta}\ln\left(\frac{4n\mu_0}{\varepsilon}\right)$, Algorithm 2 ends up with a point y such that $z^* - b^T y \leq \varepsilon$.

Complexity analysis

Maximum outer iterations

Let z^* be the optimum value of the dual problem of (P). Then, after at most $\frac{1}{\theta}\ln\left(\frac{4n\mu_0}{\varepsilon}\right)$, Algorithm 2 ends up with a point y such that $z^* - b^T y \leq \varepsilon$.

Maximum inner iterations

Each outer iteration of Algorithm 2 requires at most

$$\frac{11\theta}{(1-\theta)^2}\left(\theta n + \frac{3}{2}\sqrt{n}\right) + \frac{11}{3}$$

inner iterations.

Complexity analysis

Total number of Newton iterations

An upper bound of the total number of Newton iterations is given by

$$\left[\frac{11}{(1-\theta)^2}\left(\theta n + \frac{3}{2}\sqrt{n}\right) + \frac{11}{3}\right] \ln\left(\frac{4n\mu_0}{\varepsilon}\right)$$

Complexity analysis

Total number of Newton iterations

An upper bound of the total number of Newton iterations is given by

$$\left[\frac{11}{(1-\theta)^2}\left(\theta n + \frac{3}{2}\sqrt{n}\right) + \frac{11}{3}\right] \ln\left(\frac{4n\mu_0}{\varepsilon}\right)$$

▷ To obtain an ε-optimal solution, Algorithm 2 needs $O(n \ln \frac{n\mu_0}{\varepsilon})$ Newton iterations.

Complexity analysis

Algorithm 2

1. Initialize $\mu \leftarrow \mu_0$ and $y \leftarrow y^0$
2. (Outer loop) While $n\mu > \varepsilon$

   $\mu \leftarrow (1-\theta)\mu$

   Compute $d(y;\mu) = -H^{-1}g = (AS^{-2}A^T)^{-1}\left(\frac{b}{\mu} - AS^{-1}e\right)$

   2.1 (Inner loop) While $\|d(y;\mu)\|_H > \frac{1}{2}$

   Compute $d(y;\mu) = -H^{-1}g = (AS^{-2}A^T)^{-1}\left(\frac{b}{\mu} - AS^{-1}e\right)$

   Compute $\alpha = \text{argmin}\left\{\psi_B(y+\alpha d;\mu) : y + \alpha d \in \mathring{\mathcal{F}}(D)\right\}$

   Update $y \leftarrow y + \alpha d$

3. Output $y(\mu)$ is an ε-optimal point.

Complexity analysis

In log barrier algorithm, there are three main difficulties

- Inner loop increases the computational cost.
- Hessian matrix H suffers from ill-conditioning as $\mu \to 0$.
- Identification of α in the inner loop.

Complexity analysis

In log barrier algorithm, there are three main difficulties

- Inner loop increases the computational cost.
- Hessian matrix H suffers from ill-conditioning as $\mu \to 0$.
- Identification of α in the inner loop.

Idea

- Solve the Newton scheme directly instead of computing H^{-1}.
- Device a method that will not need the inner loop.

Contents

Primal-dual interior point method

Given LPP

$$
\text{(P)} \quad \begin{cases} \min & c^T x \\ \text{subject to} & Ax = b, x \geq 0. \end{cases}
$$

The primal-dual optimality system (perturbed KKT system) is

$$
\text{(PD OCS)} \quad \begin{cases} A^T y + s = c, \\ Ax = b, \\ Sx = \mu e, \end{cases}
$$

i.e.,

$$
F_\mu(x, y, s) \equiv \begin{bmatrix} A^T y + s - c \\ Ax - b \\ SXe - \mu e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.
$$

Primal-dual interior point method

Newton scheme to solve $F_\mu(x, y, s) = 0$ is

$$JF_\mu(x^k, y^k, s^k) \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = -F_\mu(x^k, y^k, s^k)$$

i.e., $$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} A^T y + s - c \\ Ax - b \\ SXe - \mu e \end{bmatrix}, \qquad (1)$$

where $\Delta x = x^k - x^{k+1}$, $\Delta y = y^k - y^{k+1}$ and $\Delta s = s^k - s^{k+1}$.

Primal-dual interior point method

If $(x, y, s) \in \mathring{\mathcal{F}}(P) \times \mathring{\mathcal{F}}(D)$, then the Newton system (1) reduces to

$$\text{(PD NS)} \quad \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ SXe - \mu e \end{bmatrix}$$

- Corresponding to $\mu = 0$, the system $F_\mu(x, y, s) = 0$, gives the optimum solution.
- To solve (PD NS) approximately, replace μ by $\sigma\tau(x, s)$, where σ is a regulating (centering) parameter that regulates the trade-off between the movements towards central path and optimum solution.

Primal-dual interior point method

Solve the approximate Newton system (1) reduces to

$$\text{(PD ANS)} \quad \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ SXe - \sigma\tau(x,s)e \end{bmatrix},$$

where $\tau(x,s) = \frac{1}{n}\sum_{i=1}^{n} x_i s_i$.

Primal-dual interior point method

Solve the approximate Newton system (1) reduces to

$$\text{(PD ANS)} \quad \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ SXe - \sigma\tau(x,s)e \end{bmatrix},$$

where $\tau(x,s) = \frac{1}{n}\sum_{i=1}^{n} x_i s_i$.

Observation

- If we are on the central path, then $\tau(x,s) = \mu$, but τ is also defined off the central path.
- Recall that $x_i s_i > 0$. Thus, $XSe \in \mathbb{R}^n_{++}$.The mapping $(x,y,s) \mapsto XSe$ from $\mathring{\mathcal{F}}(P) \times \mathring{\mathcal{F}}(D)$ to $\mathbb{R}^n_{++}$ is a bijection (H.W.). Hence we can study the central path in the xs-space.

Neignborhood of the central path

- Denote $\mathring{\mathcal{F}} = \mathring{\mathcal{F}}(P) \times \mathring{\mathcal{F}}(D)$

Two-norm neighborhood ($\beta \geq 0$)

$$\mathcal{N}_2(\beta) = \left\{(x, y, s) \in \mathring{\mathcal{F}} : \|XSe - \tau(x, s)e\|_2 \leq \beta\tau(x, s)\right\}.$$

One sided infinity-norm neighborhood ($0 \leq \gamma \leq 1$)

$$\mathcal{N}_{-\infty}(\gamma) = \left\{(x, y, s) \in \mathring{\mathcal{F}} : x_i s_i \geq \gamma\tau(x, s) \text{ for all } i = 1, 2, \ldots, n\right\}.$$

Neignborhood of the central path

Observation (H. W.)

- For $0 \le \beta_1 \le \beta_2 \le 1$,

$$\text{CP} = \mathcal{N}_2(0) \subset \mathcal{N}_2(\beta_1) \subset \mathcal{N}_2(\beta_2) \subset \overset{\circ}{\mathcal{F}}.$$

- For $0 \le \gamma_1 \le \gamma_2 \le 1$,

$$\text{CP} = \mathcal{N}_{-\infty}(1) \subset \mathcal{N}_{-\infty}(\gamma_2) \subset \mathcal{N}_{-}\infty(\gamma_1) \subset \mathcal{N}_{-}\infty(0) = \overset{\circ}{\mathcal{F}}.$$

Approximate Newton system at (x^k, y^k, s^k)

$$(\text{PD ANS}^k) \quad \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ S^k X^k e - \sigma \tau^k e \end{bmatrix},$$

where $\tau^k = \tau(x^k, s^k) = \frac{1}{n} \sum_{i=1}^n x_i^k s_i^k$.

Short-step path following algorithm

SPF Algorithm

Input:

- The initial starting point $(x^0, y^0, s^0) \in \mathcal{N}_2\left(\frac{1}{2}\right)$
- Gap precision $\varepsilon > 0$
- Centering parameter $\sigma = 1 - \frac{2}{5\sqrt{n}}$

1. Initialize $k \leftarrow 0$, $\tau^0 = \frac{1}{n}\sum_{i=1}^{n} x_i^0 s_i^0$
2. While $\tau^k > \varepsilon$
   - 2.1 Determine Δx^k, Δy^k and Δs^k by solving (PD ANS^k)
   - 2.2 Update $x^{k+1} = x^k - \Delta x^k$, $y^{k+1} = y^k - \Delta y^k$, $s^{k+1} = s^k - \Delta s^k$
   - 2.3 Update $\tau^k = \frac{1}{n}(x^{k+1})^T s^{k+1}$ and $k \leftarrow k + 1$.
3. Output: (x^k, y^k, s^k) is an ε-optimal point to primal and dual.

Complexity of SPF method

Result for SPF method

- Let $(x^0, y^0, s^0) \in \mathcal{N}_2\left(\frac{1}{2}\right)$ and $\tau_0 < \frac{1}{\varepsilon^\nu}$ for some $\nu > 0$ and a given $\varepsilon \in (0,1)$. Then, there exists an index $K \in O(\sqrt{n}|\ln \varepsilon|)$ such that

$$\tau^k < \varepsilon \text{ for all } k \geq K$$

Complexity of SPF method

Result for SPF method

- Let $(x^0, y^0, s^0) \in \mathcal{N}_2\left(\frac{1}{2}\right)$ and $\tau_0 < \frac{1}{\varepsilon^\nu}$ for some $\nu > 0$ and a given $\varepsilon \in (0,1)$. Then, there exists an index $K \in O(\sqrt{n}|\ln \varepsilon|)$ such that

$$\tau^k < \varepsilon \text{ for all } k \geq K$$

- Let the parameters $\beta \in (0,1)$ and $\sigma \in (0,1)$ be chosen to satisfy

$$\frac{\beta^2 + n(1-\sigma)^2}{\sqrt{8}(1-\beta)} \leq \sigma\beta.$$

  If $(x^k, y^k, s^k) \in \mathcal{N}_2(\beta)$, we have

$$(x^{k+1}, y^{k+1}, s^{k+1}) \in \mathcal{N}_2(\beta).$$

Long-step path following algorithm

LPF Algorithm

Input:

- The initial starting point $(x^0, y^0, s^0) \in \mathcal{N}_{-\infty}(\gamma)$
- Gap precision $\varepsilon > 0$
- $\sigma_{\min}$ and $\sigma_{\max}$ with $0 < \sigma_{\min} < \sigma_{\max} < 1$.
- The neighborhood parameter $\gamma \in (0, 1)$.
- Initialize: $k \leftarrow 0$, $\tau^0 = \frac{1}{n} \sum_{i=1}^{n} x_i^0 s_i^0$.

Long-step path following algorithm

LPF Algorithm

1. While $\tau^k > \varepsilon$
   1.1 Determine Δx^k, Δy^k and Δs^k by solving (PD ANS^k) for a $\sigma^k \in [\sigma_{\min}, \sigma_{\max}]$
   1.2 Update $x^{k+1} = x^k - \Delta x^k$, $y^{k+1} = y^k - \Delta y^k$, $s^{k+1} = s^k - \Delta s^k$
   1.3 Find the largest $\alpha_k \in (0, 1]$ such that
   $$(x^k - \alpha_k \Delta y^k, y^k - \alpha_k \Delta y^k, s^k - \alpha_k \Delta s^k) \in \mathcal{N}_{-\infty}(\gamma)$$
   1.4 Update
   $x^{k+1} = x^k - \alpha_k \Delta y^k, y^{k+1} = y^k - \alpha_k \Delta y^k, s^{k+1} = s^k - \alpha_k \Delta s^k$
   1.5 Update $\tau^k = \frac{1}{n}(x^{k+1})^T s^{k+1}$ and $k \leftarrow k + 1$.
2. Output: (x^k, y^k, s^k) is an ε-optimal point to primal and dual.

Complexity of LPF method

Result for LPF method

- Let $(x^0, y^0, s^0) \in \mathcal{N}_{-\infty}(\gamma)$ and $\tau_0 < \frac{1}{\varepsilon^\nu}$ for some $\nu > 0$ and a given $\varepsilon \in (0, 1)$. Then, there exists an index $K \in O(n|\ln \varepsilon|)$ such that

$$\tau^k < \varepsilon \text{ for all } k \geq K$$

References

1. D. den Hertog (2012). *Interior point approach to linear, quadratic and convex programming: algorithms and complexity*, Springer.
2. W. Forst and D. Hoffmann (2010). *Optimization–theory and practice*. Springer.
3. A. V. Fiacco and G. P. McCormick (1987). *Nonlinear programming: sequential unconstrained minimization techniques*, Classics in applied mathematics, Vol. 4, Society for Industrial and Applied Mathematics.
4. R. J. Vanderbei (2020). *Linear programming: foundations and extensions*. Vol. 285. Springer Nature, 2020.
5. Wright, M. (2005). The interior-point revolution in optimization: history, recent developments, and lasting consequences, *Bulletin of the American mathematical society* 42(1), 39–56.