Primal-dual interior-point methods part I

Javier Peña (guest lecturer)
Convex Optimization 10-725/36-725

Last time: barrier method

Approximate

$$\begin{array}{ll} \min & f(x) \\ \text{subject to} & h_i(x) \le 0, \quad i = 1, \ldots m \\ & Ax = b \end{array}$$

with

$$\begin{array}{ll} \min & tf(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

where ϕ is the log-barrier function

$$\phi(x) = -\sum_{i=1}^{m} \log(-h_i(x)).$$

Barrier method

Solve a sequence of problems

$$\begin{array}{ll} \min & tf(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

for increasing values of $t > 0$, until $m/t \le \epsilon$.

Start with $t = t^{(0)} > 0$, and solve the above problem using Newton's method to produce $x^{(0)} = x^\star(t)$.

For $k = 1, 2, 3, \ldots$

- Solve the barrier problem at $t = t^{(k)}$, using Newton's method initialized at $x^{(k-1)}$, to produce $x^{(k)} = x^\star(t)$
- Stop if $m/t \le \epsilon$
- Else update $t^{(k+1)} = \mu t$, where $\mu > 1$

Outline

Today:

- Recap of linear programming and duality
- The central path
- Feasible path-following interior-point methods
- Infeasible interior-point methods

Linear program in standard form

Problem of the form

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m, c \in \mathbb{R}^n$.

Recall: Any linear program can be rewritten in standard form.

Standard assumption: A is full row-rank.

Linear programming duality

The dual of the above problem is

$$\begin{array}{ll} \max\limits_{y} & b^{\mathsf{T}} y \\ \text{subject to} & A^{\mathsf{T}} y \leq c. \end{array}$$

Or equivalently

$$\begin{array}{ll} \max\limits_{y,s} & b^{\mathsf{T}} y \\ \text{subject to} & A^{\mathsf{T}} y + s = c \\ & s \geq 0. \end{array}$$

Throughout the sequel refer to the LP from the previous slide as the primal problem and to the above LP as the dual problem.

Linear programming duality

Theorem (Weak duality)

Assume x is primal feasible and y is dual feasible. Then

$$b^\mathsf{T} y \leq c^\mathsf{T} x.$$

Theorem (Strong duality)

Assume primal LP is feasible. Then it is bounded if and only if the dual is feasible. In that case their optimal values are the same and they are attained.

Optimality conditions

The points x^* and (y^*, s^*) are respectively primal and dual optimal solutions if and only if (x^*, y^*, s^*) solves

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

Two main classes of algorithms for linear programming

Simplex method: Maintain first three conditions and aim for the fourth one.

Interior-point methods: Maintain first two and the fourth conditions and aim for the third one.

Some history

- Dantzig (1940s): the simplex method. Still one of the most popular algorithms for linear programming.
- Klee and Minty (1960s): LP with n variables and $2n$ constraints that the simplex method needs to perform 2^n iterations to solve.
- Khachiyan (1979): first polynomial-time algorithm for LP based on the ellipsoid method of Nemirovski and Yudin (1976). Theoretically strong but computationally weak.
- Karmarkar (1984): first interior-point polynomial-time algorithm for LP.
- Renegar (1988): Newton-based interior-point algorithm for LP. Best known theoretical complexity to date.
- Modern state-of-the-art LP solvers typically use both simplex and interior-point methods.

Barrier method for primal and dual problems

Pick $\tau > 0$. Approximate the LP primal with

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x - \tau \sum_{j=1}^{n} \log x_j \\ \text{subject to} \quad & Ax = b \end{aligned}$$

and the LP dual with

$$\begin{aligned} \max_{y,s} \quad & b^\mathsf{T} y + \tau \sum_{j=1}^{n} \log s_j \\ \text{subject to} \quad & A^\mathsf{T} y + s = c. \end{aligned}$$

Neat fact:
The above two problems are, modulo a constant, Lagrangian duals of each other.

Primal-dual central path

Assume the primal and dual problems are strictly feasible. The primal-dual central path is the set

$$\{(x(\tau), y(\tau), s(\tau)) : \tau > 0\}$$

where $x(\tau)$, and $(y(\tau), s(\tau))$ solve the above pair of barrier problems. Equivalently, $(x(\tau), y(\tau), s(\tau))$ is the solution to

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

Path following interior-point methods

Main idea: Generate $(x^k, y^k, s^k) \approx (x(\tau^k), y(\tau^k), s(\tau^k))$ for $\tau^k \downarrow 0$.

Key details:

- Proximity to the central path
- Decrease τ^k
- Update (x^k, y^k, s^k)

Neighborhoods of the central path

Notation:

- $\mathcal{F}^0 := \{(x, y, s) : Ax = b,\ A^\mathsf{T} y + s = c,\ x, s > 0\}$.
- For $x, s \in \mathbb{R}^n$, $X := \mathrm{diag}(x),\ S := \mathrm{diag}(s)$.
- Given $x, s \in \mathbb{R}^n_+$, $\mu(x, s) := \frac{x^\mathsf{T} s}{n}$

For $\theta \in (0, 1)$, two-norm neighborhood:

$$\mathcal{N}_2(\theta) := \{(x, y, s) \in \mathcal{F}^0 : \|XS\mathbf{1} - \mu(x, s)\mathbf{1}\|_2 \leq \theta\mu(x, s)\}$$

For $\gamma \in (0, 1)$, one-sided infinity-norm neighborhood:

$$\mathcal{N}_{-\infty}(\gamma) := \{(x, y, s) \in \mathcal{F}^0 : x_i s_i \geq \gamma\mu(x, s),\ i = 1, \ldots, n\}$$

Newton step

Recall: $(x(\tau), y(\tau), s(\tau))$ solution to

$$\begin{bmatrix} A^\mathsf{T} y + s - c \\ Ax - b \\ XS\mathbf{1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau\mathbf{1} \end{bmatrix}, \; x, s > 0.$$

Newton step equations:

$$\begin{bmatrix} 0 & A^\mathsf{T} & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau\mathbf{1} - XS\mathbf{1} \end{bmatrix}.$$

Short-step path following algorithm

Algorithm SPF

1. Let $\theta, \delta \in (0,1)$ be such that $\frac{\theta^2+\delta^2}{2^{3/2}(1-\theta)} \leq \left(1 - \frac{\delta}{\sqrt{n}}\right)\theta$.
2. Let $(x^0, y^0, s^0) \in \mathcal{N}_2(\theta)$.
3. For $k = 0, 1, \ldots$
    - Compute Newton step for
      $$(x, y, s) = (x^k, y^k, s^k),\ \tau = \left(1 - \frac{\delta}{\sqrt{n}}\right)\mu(x, s).$$
    - Set $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + (\Delta x, \Delta y, \Delta s)$.

Theorem

The sequence generated by Algorithm SPF satisfies

$$(x^k, y^k, s^k) \in \mathcal{N}_2(\theta),$$

and

$$\mu(x^{k+1}, s^{k+1}) = \left(1 - \frac{\delta}{\sqrt{n}}\right) \mu(x^k, s^k)$$

Corollary

In $\mathcal{O}\left(\sqrt{n}\log\left(\frac{n\mu(x^0,s^0)}{\epsilon}\right)\right)$ *the algorithm yields* $(x^k, y^k, s^k) \in \mathcal{F}^0$ *such that*

$$c^\mathsf{T} x_k - b^\mathsf{T} y_k \leq \epsilon.$$

Long-step path following algorithm

Algorithm LPF

1. Choose $\gamma \in (0,1)$ and $0 < \sigma_{\min} < \sigma_{\max} < 1$
2. Let $(x^0, y^0, s^0) \in \mathcal{N}_{-\infty}(\gamma)$
3. For $k = 0, 1, \ldots$
   - Choose $\sigma \in [\sigma_{\min}, \sigma_{\max}]$
   - Compute Newton step for
     $$(x, y, s) = (x^k, y^k, s^k),\ \tau = \sigma\mu(x^k, s^k)$$
   - Choose α_k as the largest $\alpha \in [0, 1]$ such that
     $$(x^k, y^k, s^k) + \alpha(\Delta x, \Delta y, \Delta s) \in \mathcal{N}_{-\infty}(\gamma)$$
   - Set $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + \alpha_k(\Delta x, \Delta y, \Delta s)$

Theorem

The sequence generated by Algorithm LPF satisfies

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

and

$$\mu(x^{k+1}, s^{k+1}) \leq \left(1 - \frac{\delta}{n}\right) \mu(x^k, s^k)$$

for some constant δ *that depends on* $\gamma, \sigma_{\min}, \sigma_{\max}$ *but not on* n.

Corollary

In $\mathcal{O}\left(n \log\left(\frac{n\mu(x^0,s^0)}{\epsilon}\right)\right)$ *the algorithm yields* $(x^k, y^k, s^k) \in \mathcal{F}^0$ *such that*

$$c^\mathsf{T} x_k - b^\mathsf{T} y_k \leq \epsilon.$$

Infeasible interior-point algorithms

Algorithms SPF and LPF require an initial point in $\mathcal{F}^0$.

Can we eliminate this requirement?

Given (x, y, s), let $r_b := Ax - b, r_c := A^\mathsf{T} y + s - c$.

Assume (x^0, y^0, s^0) with $x^0, s^0 > 0$ is given. Extend $\mathcal{N}_{-\infty}(\gamma)$ to

$$\begin{aligned}\mathcal{N}_{-\infty}(\gamma, \beta) := \{(x, y, s) : &\|(r_b, r_c)\| \le [\|(r_b^0, r_c^0)\|/\mu^0]\beta\mu, \\ &x, s > 0, x_i s_i \ge \gamma\mu, i = 1, \ldots, n\}\end{aligned}$$

for $\gamma \in (0, 1)$, $\beta \ge 1$.

Newton step equations:

$$\begin{bmatrix} 0 & A^\mathsf{T} & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} r_c \\ r_b \\ XS\mathbf{1} - \tau\mathbf{1} \end{bmatrix}.$$

Algorithm IPF

1. Choose $\gamma \in (0,1)$, $0 < \sigma_{\min} < \sigma_{\max} < 0.5$, and $\beta \geq 1$.
2. Choose (x^0, y^0, s^0) with $x^0, s^0 > 0$
3. For $k = 0, 1, \ldots$
    - Choose $\sigma \in [\sigma_{\min}, \sigma_{\max}]$
    - Compute Newton step for
    $$(x, y, s) = (x^k, y^k, s^k),\ \tau = \sigma\mu(x^k, s^k)$$
    - Choose α_k as the largest $\alpha \in [0,1]$ such that
    $$(x^k, y^k, s^k) + \alpha(\Delta x, \Delta y, \Delta s) \in \mathcal{N}_{-\infty}(\gamma, \beta)$$
    and
    $$\mu(x^k + \alpha\Delta x, s^k + \alpha\Delta s) \leq (1 - 0.01\alpha)\mu(x_k, s_k)$$
    - Set $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + \alpha_k(\Delta x, \Delta y, \Delta s)$

Theorem

Assume the primal and dual problems have optimal solutions. Then the sequence $(x^k, y^k, s^k), k = 0, 1, \ldots$ *generated by Algorithm IPF satisfies*

$$\mu_k := \mu(x^k, y^k, s^k) \to 0 \ \text{ linearly.}$$

In particular

$$\|(r_b^k, r_c^k)\| \to 0 \ \text{ R-linearly.}$$

Remark: $a_k \to 0$ R-linearly $\Leftrightarrow |a_k| \leq b_k$ and $b_k \to 0$ linearly.

IPM for more general convex optimization

Consider a convex minimization problem

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & h(x) \le 0 \\ & Ax = b. \end{aligned}$$

Assume f, h smooth and strong duality holds. Then x^* and (u^*, v^*) are respectively primal and dual optimal solutions if and only if (x^*, u^*, v^*) solves the KKT conditions

$$\begin{aligned} \nabla f(x) + A^\mathsf{T} v + \nabla h(x) u &= 0 \\ U h(x) &= 0 \\ Ax &= b \\ u, -h(x) &\ge 0. \end{aligned}$$

Central path and Newton step

Central path:
$\{(x(\tau), u(\tau), v(\tau)) : \tau > 0\}$ where $(x(\tau), u(\tau), v(\tau))$ solves

$$\begin{aligned} \nabla f(x) + A^\mathsf{T} v + \nabla h(x) u &= 0 \\ U h(x) &= -\tau \mathbf{1} \\ Ax &= b \\ u, -h(x) &> 0. \end{aligned}$$

Newton step:

$$\begin{bmatrix} \nabla^2 f(x) + \sum_i u_i \nabla^2 h_i(x) & \nabla h(x) & A^\mathsf{T} \\ U \nabla h(x)^\mathsf{T} & H(x) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \\ \Delta v \end{bmatrix} = - \begin{bmatrix} r_{\mathsf{dual}} \\ r_{\mathsf{cent}} \\ r_{\mathsf{pri}} \end{bmatrix},$$

$r_{\mathsf{dual}} = \nabla f(x) + A^\mathsf{T} v + \nabla h(x) u,\ r_{\mathsf{cent}} = U h(x) + \tau \mathbf{1},\ r_{\mathsf{pri}} = Ax - b.$

Given x, u such that $h(x) \leq 0,\ u \geq 0$, define $\mu(x,u) := -\frac{h(x)^\mathsf{T} u}{m}$.

Primal-Dual Algorithm

1. Choose $\sigma \in (0,1)$
2. Choose (x^0, u^0, v^0) such that $h(x^0) < 0,\ u^0 > 0$
3. For $k = 0, 1, \ldots$
   - Compute Newton step for

$$(x,u,v) = (x^k, u^k, v^k),\ \tau := \sigma\mu(x^k, u^k)$$

   - Choose steplength α_k via line-search and set

$$(x^{k+1}, u^{k+1}, v^{k+1}) := (x^k, u^k, v^k) + \alpha_k(\Delta x, \Delta u, \Delta v)$$

Line-search:
Maintain $h(x) < 0,\ u > 0$ and reduce $\|r_{\text{dual}}\|, \|r_{\text{cent}}\|, \|r_{\text{pri}}\|$.

References and further reading

- S. Wright (1997), "Primal-Dual Interior-Point Methods," Chapters 5 and 6
- S. Boyd and L. Vandenberghe (2004), "Convex optimization," Chapter 11
- J. Renegar (2001), "A Mathematical View of Interior-Point Methods"