Lecture 7: Interior Point Methods for Constrained Minimization

- brief intro of IP methods
- logarithmic barrier function
- central path
- SUMT
- feasibility phase
- summary

Introduction of IP methods

interior point methods

- smooth 'barrier' function replaces constraints
- solve sequence of smooth unconstrained problems
- initiated by Karmarkar (for LP)
- polynomial worst-case complexity
- work well in practice
- extended to general case by Nesterov & Nemirovsky 1988

Logarithmic barrier function

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m\end{array}$$

f_i convex, differentiable; (no equality constraints for simplicity)

assume strict feasibility: $C = \{x \mid f_i(x) < 0,\ i = 1, \ldots, m\} \neq \emptyset$

define **logarithmic barrier** ϕ as

$$\phi(x) = \begin{cases} -\sum_{i=1}^{m} \log(-f_i(x)) & x \in C \\ +\infty & \text{otherwise} \end{cases}$$

- ϕ is convex, smooth on C
- $\phi \to \infty$ as x approaches boundary of C

$\operatorname{argmin} \phi$ (if exists) is called **analytic center** of inequalities $f_1(x) < 0, \ldots, f_m(x) < 0$

Central path

$$x^*(t) = \operatorname{argmin}(tf_0(x) + \phi(x)) \text{ for } t > 0$$

(we assume minimizer exists and is unique)

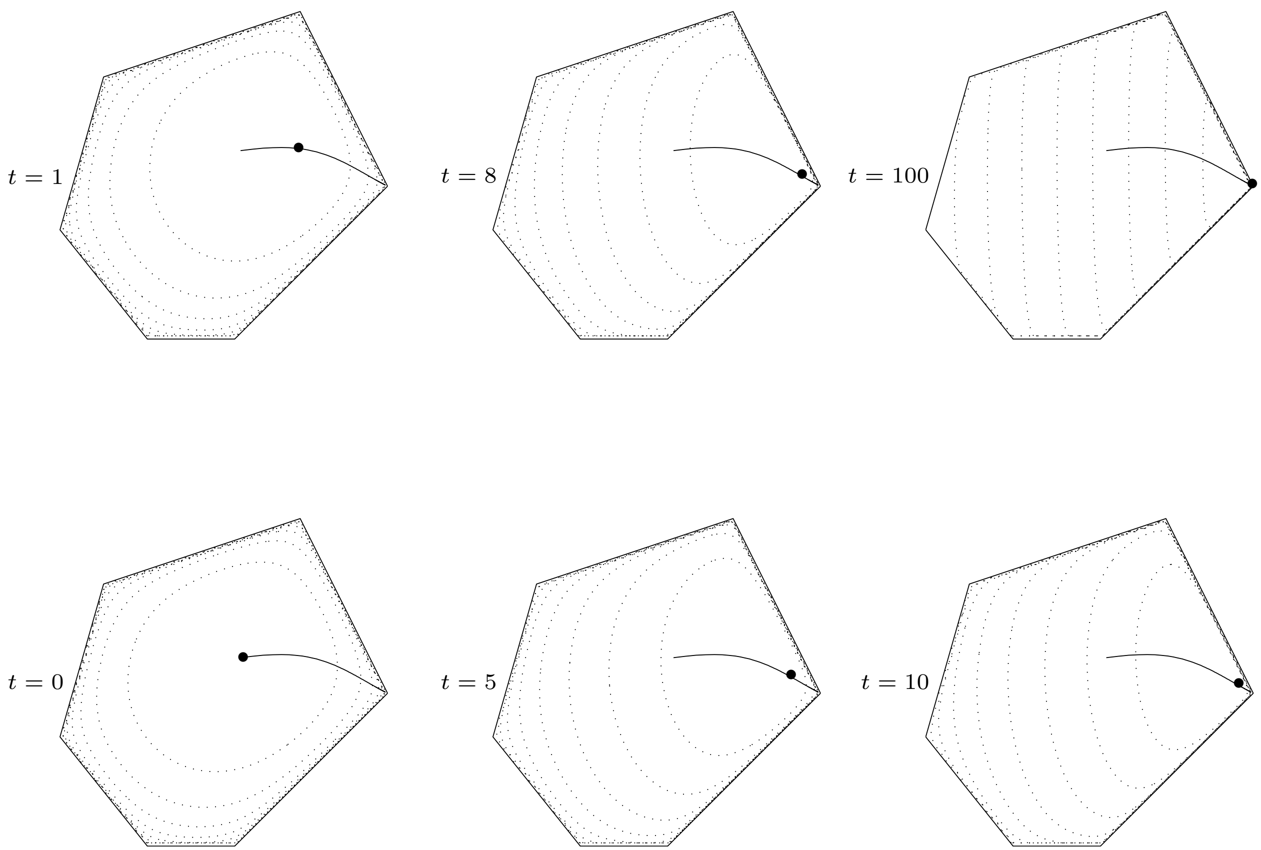
- curve $x^*(t)$ for $t \geq 0$ called **central path**
- can compute $x^*(t)$ by solving smooth unconstrained minimization problem (given a strictly feasible starting point)
- t gives relative weight of objective and barrier
- barrier 'traps' $x^*(t)$ in strictly feasible set
- intuition suggests $x^*(t)$ converges to optimal as $t \to \infty$

$x^*(t)$ characterized by

$$t\nabla f_0(x^*(t)) + \sum_{i=1}^{m} \frac{1}{-f_i(x^*(t))} \nabla f_i(x^*(t)) = 0$$

Example: central path for LP

$x \in \mathbf{R}^2$, $A \in \mathbf{R}^{6\times 2}$, c points left



Force field interpretation

imagine a particle in C, subject to forces; ith constraint generates **constraint force field**

$$F_i(x) = -\nabla\left(-\log(-f_i(x))\right) = \frac{1}{f_i(x)}\nabla f_i(x)$$

- ϕ is *potential* associated with constraint forces
- constraint forces push particle away from boundary of feasible set
- constraint forces trap particle in C

superimpose **objective force field**

$$F_0(x) = -t\nabla f_0(x)$$

- pulls particle toward small f_0
- t scales objective force

at $x^*(t)$, constraint forces balance objective force;

as t increases, particle is pulled towards optimal point, trapped in C by barrier potential

Central points and duality

recall $x^* = x^*(t)$ satisfies

$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x^*) = 0, \quad \lambda_i = \frac{1}{-f_i(x^*)t} > 0$$

so x^* also minimizes $L(x, \lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$

$i.e.$, λ is dual feasible and

$$f^\star \geq g(\lambda) = \inf_x \left(f_0(x) + \sum_i \lambda_i f_i(x) \right) = f_0(x^*) + \sum_i \lambda_i f_i(x^*) = f_0(x^*) - m/t$$

summary: a point on central path yields dual feasible point and lower bound:

$$f_0(x^*(t)) \geq p^\star \geq f_0(x^*(t)) - m/t$$

(which proves $x^*(t)$ becomes optimal as $t \to \infty$)

Central path and KKT conditions

KKT optimality conditions: x optimal $\iff \exists\lambda$ s.t.

$$\begin{array}{rcl} f_i(x) & \leq & 0 \\ \lambda_i & \geq & 0 \\ \nabla f_0(x) + \sum_i \lambda_i \nabla f_i(x) & = & 0 \\ \lambda_i f_i(x) & = & 0 \end{array}$$

centrality conditions: x central $\iff \exists\lambda$, $t > 0$ s.t.

$$\begin{array}{rcl} f_i(x) & \leq & 0 \\ \lambda_i & \geq & 0 \\ \nabla f_0(x) + \sum_i \lambda_i \nabla f_i(x) & = & 0 \\ \lambda_i f_i(x) & = & -1/t \end{array}$$

- for t large, $x^*(t)$ 'almost' satisfies KKT
- central path is continuous deformation of KKT condition

Unconstrained minimization method

given strictly feasible x, desired accuracy $\epsilon > 0$

1. $t := m/\epsilon$
2. compute $x^*(t)$ starting from x
3. $x := x^*(t)$

- computes ϵ-suboptimal point on central path
  (and certificate λ)
- solves constrained problem by solving one smooth unconstrained minimization
  (via Newton, BFGS, . . .)
- works, but can be slow

SUMT

(Sequential Unconstrained Minimization Technique)

given strictly feasible x, $t > 0$, tolerance $\epsilon > 0$
repeat
1. compute $x^*(t)$ starting from x
2. $x := x^*(t)$
3. if $m/t \leq \epsilon$, return(x)
4. increase t

- generates sequence of points on central path
- solves constrained problem via sequence of unconstrained minimizations (often, Newton)
- simple updating rule for t: $t^+ = \mu t$
  (typical values $\mu \approx 10 \sim 100$)

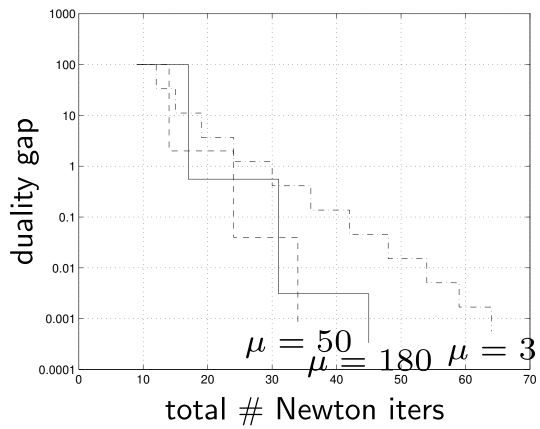
steps 1–4 above called **outer iteration** step 1 involves **inner iterations** ($e.g.$, Newton)

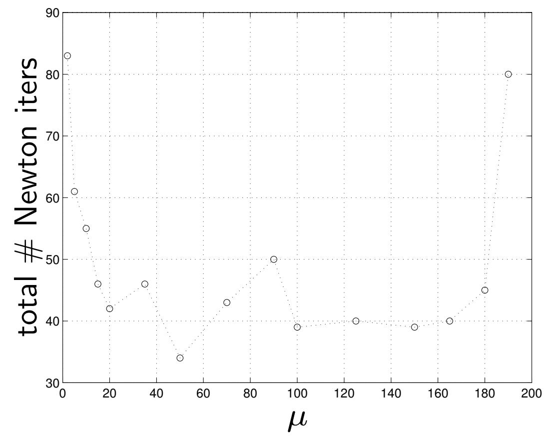
tradeoff: small $\mu \Longrightarrow$ few inner iters to compute $x^{(k+1)}$ from $x^{(k)}$, but more outer iters

Example: LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

$A \in \mathbf{R}^{100\times 50}$, Newton with exact line search



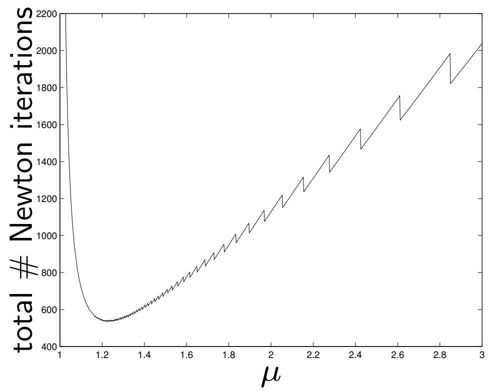


- (Left): duality gap reduced by 10^5 in few tens of Newton iters;
- (Right): trade-off in choice of μ: #Newton iters required to reduce duality gap by 10^6

Bound on total # Newton iters

upper bound on total #Newton steps:

$$\left\lceil \frac{\log(m/t^{(0)}\epsilon)}{\log \mu} \right\rceil \left(c + \frac{m(\mu - 1 - \log \mu)}{\eta_2} \right)$$



($c = 5$, $\eta_2 = 1/20$, $m = 10$, $m/t^{(0)}\epsilon = 10^5$)

- confirms trade-off in choice of μ
- optimal μ depends on m, η_2, c, $t^{(0)}$, ϵ

could use empirical values for η_2, c to optimize average-case behavior

Phase I

to compute strictly feasible point (or determine none exists) set up auxiliary problem:

$$\begin{array}{ll} \text{minimize} & w \\ \text{subject to} & f_i(x) \le w, \;\; i = 1, \ldots, m \end{array}$$

- easy to find strictly feasible initial point (hence SUMT can be used)
- can use stopping criterion with target value 0

Generalized inequalities

standard problem with generalized inequalities:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, \;\; i = 1, \ldots, L \end{array}$$

- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is convex, differentiable
- $f_i : \mathbf{R}^n \to \mathbf{R}^{m_i}$ are K_i-convex, differentiable

ψ is a **log barrier** for cone $K \subseteq \mathbf{R}^m$ if

- $\mathbf{dom}\,\psi = \mathbf{int}\,K$
- ψ is convex and K-increasing
- there is a θ s.t. for all $a > 0$, $z \succ_K 0$,

$$\psi(az) = \psi(z) - \theta \log a$$

generalizes logarithm from $\mathbf{R}_+$ to cone K

example. $\psi(Z) = \log\det Z^{-1}$ is a log barrier for PSD cone $K \subseteq \mathbf{R}^{n\times n}$, with $\theta = n$

Minimization with equality constraint

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- Newton step for min $f(x) = tf_0(x) + \phi(x)$: min 2nd order expansion of $f(x + \Delta x)$,

$$\begin{array}{ll} \text{minimize} & f(x) + \nabla f(x)^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(x) \Delta x \\ \text{subject to} & A\Delta x = 0 \end{array}$$

- solves KKT conditions:

$$0 = \nabla f(x) + \nabla^2 f(x) \Delta x + A^T w; \; A\Delta x = 0$$

- equality constrained Newton step Δx:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

Summary

- other IP methods similar to SUMT
- work very well in practice
- worst-case complexity theory (if self-concordant)

sophisticated variations:

- use predictor steps to follow central path, with aggressive step size rules ($e.g.$, 99% to boundary!)
- primal-dual methods
- infeasible methods (combine phase I & II)
- incomplete centering