Convex Optimization

Lecture 13 - Interior-Point Methods

Instructor: Yuanzhang Xiao

University of Hawaii at Manoa

Fall 2017

Today's Lecture

1. Basic Concepts
2. The Barrier Method
3. Phase I Method For Infeasible Start
4. Problems With Generalized Inequalities
5. Primal-Dual Interior-Point Methods
6. Implementation Issues

Outline

Inequality Constrained Optimization Problems

inequality constrained minimization problem:

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

- $f_0, f_1, \ldots, f_m$ convex, twice continuously differentiable
- $A \in \mathbb{R}^{p \times n}$ with $\textbf{rank} A = p < n$
- $p^\star = f_0(x^\star)$ attained and finite
- problem is strictly feasible: there exists x such that

$$x \in \textbf{dom} f_0, \quad f_i(x) < 0, \;\; i = 1, \ldots, m, \quad Ax = b$$

strong duality holds, KKT conditions are sufficient and necessary

"Hierarchy" of Convex Optimization Algorithms

equality constrained quadratic problem:

- solve a linear system analytically or in one shot

equality constrained general problem:

- Newton's method
- a series of equality constrained quadratic problems

inequality constrained general problem:

- interior-point methods
- a series of equality constrained general problems

Equivalent Reformulation Using Indicator Function

equivalent reformulation using indicator functions:

$$\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{i=1}^{m} I_{-}(f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

where $I_{-} : \mathbb{R} \to \mathbb{R}$ is the indicator function

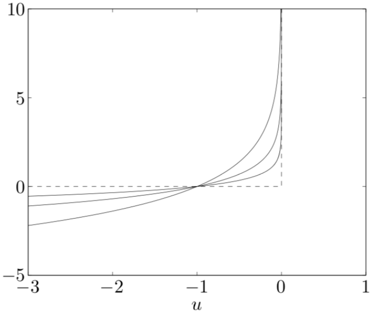
$$I_{-}(u) = \begin{cases} 0, & u \leq 0 \\ \infty, & u > 0 \end{cases}$$

equality constrained problem with undesirable objective functions

Logarithmic Barrier Function

approximate the indicator function by:

$$\widehat{I}_-(u) = -(1/t)\log(-u), \quad \mathbf{dom}\,\widehat{I}_- = -\mathbb{R}_{++}$$



approximation is more accurate as $t \to \infty$

Logarithmic Barrier Function

equality constrained problem with nice objective functions

$$\begin{array}{ll} \text{minimize} & f_0(x) - (1/t)\sum_{i=1}^{m} \log(-f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

logarithmic barrier function:

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \quad \textbf{dom}\,\phi = \{x \mid f_i(x) < 0,\ i = 1, \ldots, m\}$$

- convex and twice continuously differentiable

$$\begin{aligned} \nabla \phi(x) &= \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) \\ \nabla^2 \phi(x) &= \sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla^2 f_i(x) \end{aligned}$$

Central Path

consider the problem

$$\begin{array}{ll} \text{minimize} & tf_0(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

- denote the solution by $x^\star(t)$
- central path: a sequence of points $x^\star(t)$ as $t \to \infty$

properties of points on the central path:

- strictly feasible:

$$f_i(x^\star(t)) < 0, \;\; i = 1, \ldots, m, \quad Ax^\star(t) = b$$

- KKT condition: there exists $\hat{\nu} \in \mathbb{R}^p$ such that

$$\begin{aligned} 0 &= t\nabla f_0(x^\star(t)) + \nabla\phi(x^\star(t)) + A^T\hat{\nu} \\ &= t\nabla f_0(x^\star(t)) + \sum_{i=1}^m \frac{1}{-f_i(x^\star(t))}\nabla f_i(x^\star(t)) + A^T\hat{\nu} \end{aligned}$$

Illustration of Central Path

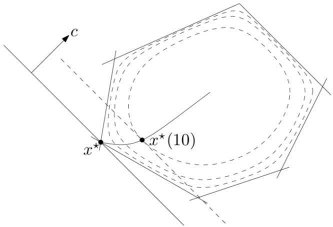
when there is no equality constraint:

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

at $x^\star(t)$, gradient of f_0 is parallel to gradient of ϕ

LP in inequality form

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



Dual Points on Central Path

recall:

$$0 = \nabla f_0(x^\star(t)) + (1/t)\sum_{i=1}^m \frac{1}{-f_i(x^\star(t))}\nabla f_i(x^\star(t)) + (1/t)A^T\hat{\nu}$$

define dual variables

$$\lambda_i^\star(t) = \frac{1}{-tf_i(x^\star(t))}, \quad i = 1,\ldots,m, \quad \nu^\star(t) = \hat{\nu}/t$$

$x^\star(t)$ minimizes Lagrangian $L(x,\lambda,\nu)$ at $\lambda_i^\star(t), \nu^\star(t)$, because

$$0 = \nabla f_0(x^\star(t)) + \sum_{i=1}^m \lambda_i^\star(t)\nabla f_i(x^\star(t)) + A^T\nu^\star(t)$$

dual function

$$\begin{aligned} & g(\lambda_i^\star(t), \nu^\star(t)) \\ = \; & f_0(x^\star(t)) + \sum_{i=1}^m \lambda_i^\star(t) f_i(x^\star(t)) + \nu^\star(t)^T(Ax^\star(t) - b) \\ = \; & f_0(x^\star(t)) - m/t \end{aligned}$$

we have $f_0(x^\star(t)) - p^\star \leq= m/t$

Interpretation Via KKT Conditions

$x^\star(t)$ and $\lambda^\star(t)$, $\nu^\star(t)$ defined above satisfy:

$$
\begin{aligned}
Ax^\star(t) = b,\ f_i(x^\star(t)) &\leq 0,\ i = 1, \ldots, m \\
\lambda^\star(t) &\geq 0 \\
\nabla f_0(x^\star(t)) + \textstyle\sum_{i=1}^m \lambda_i^\star(t) \nabla f_i(x^\star(t)) + A^T \nu^\star(t) &= 0 \\
\lambda_i^\star(t) f_i(x^\star(t)) &= -1/t,\ i = 1, \ldots, m
\end{aligned}
$$

"almost" satisfy KKT conditions, except complementary slackness

as $t \to \infty$, $x^\star(t)$, $\lambda^\star(t)$, $\nu^\star(t)$ satisfy KKT conditions

Outline

The Barrier Method

the barrier method with strictly feasible starting point:

- strictly feasible starting point x, $t > 0$, $\mu > 1$, tolerance $\epsilon > 0$
- repeat the following steps
  1. centering step:

     compute $x^\star(t)$ by minimizing $tf_0 + \phi$ subject to $Ax = b$, starting at x

  2. update $x := x^\star(t)$
  3. quit if $m/t < \epsilon$
  4. increase $t := \mu t$

features:

- use $x^\star(t^k)$ as starting point for solving for. $x^\star(t^{k+1})$
- outer iterations: centering steps
- inner iterations: Newton iterations in one centering step

Implementation Issues

accuracy of centering:

- computing $x^\star(t)$ exactly or with reasonable accuracy

choice of μ

- trade-off in the numbers of inner and outer iterations

choice of initial $t^{(0)}$

- small $t^{(0)} \rightarrow$ fewer inner iterations in the first outer iteration, but more outer iterations

using infeasible start Newton method

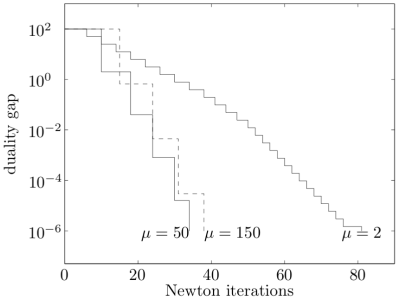
- starting point $x^{(0)}$ do not necessarily satisfy $Ax^{(0)} = b$
- still need to satisfy $f_i(x^{(0)}) < 0, \ i = 1, \ldots, m$

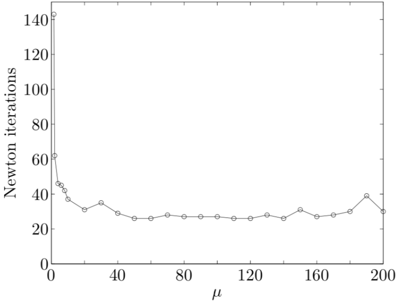
Examples – LP in Inequality Form

LP in inequality form

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

with $A \in \mathbb{R}^{100\times 50}$





Convergence Results

number of outer iterations is exactly:

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

number of inner iterations

- Newton's methods: quadratic convergence
- as t increases, the number of inner iterations nearly constant

Outline

Phase I Methods For Infeasible Start

what if we do not know which x is feasible?

- Phase I: compute a strictly feasible point
- Phase II: barrier methods

basic phase I method:

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

with optimization variables $x \in \mathbb{R}^n$ and $s \in \mathbb{R}$

- get a strictly feasible starting point (by making s large)
- use barrier methods

Basic Phase I Method

basic phase I method:

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

with optimization variables $x \in \mathbb{R}^n$ and $s \in \mathbb{R}$

suppose that the optimal value is $\bar{p}^\star$

- $\bar{p}^\star < 0$: exists a strictly feasible point
  - strictly feasible x found in the process (early termination)
- $\bar{p}^\star > 0$: original problem is infeasible
- $\bar{p}^\star = 0$: exists no strictly feasible point

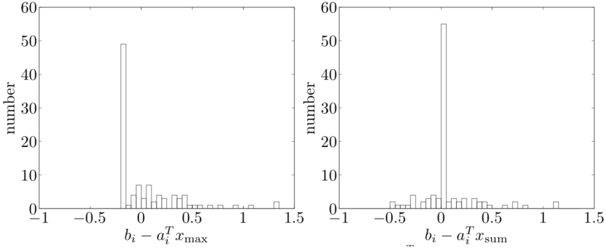
Sum of Infeasibilities Phase I Method

sum of infeasibilities phase I method:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T s \\ \text{subject to} & f_i(x) \leq s_i, \; i = 1, \ldots, m \\ & s \geq 0 \\ & Ax = b \end{array}$$

with optimization variables $x \in \mathbb{R}^n$ and $s \in \mathbb{R}^m$

for infeasible problems, finds solutions satisfying more inequalities



Feasibility Via Infeasible Start Newton Method

equivalent problem:

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

use infeasible start Newton method to solve

$$\begin{array}{ll} \text{minimize} & tf_0(x) - \sum_{i=1}^{m} \log\left(s - f_i(x)\right) \\ \text{subject to} & Ax = b, \ \ s = 0 \end{array}$$

initialize with a starting point (x, s) that :

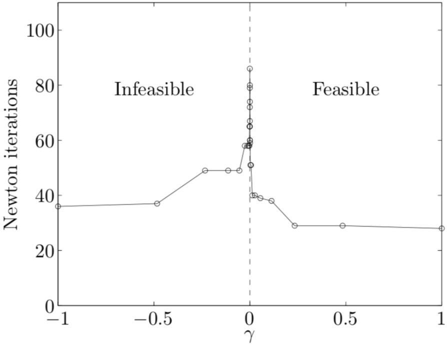
- satisfies $f_i(x) < s, \ \ i = 1, \ldots, m$
- not necessarily satisfies $Ax = b$ or $s = 0$

Example - Phase I Method

linear feasibility problems:

$$Ax \leq b + \gamma \Delta b$$

- $A \in \mathbb{R}^{50 \times 20}$
- strictly feasible for $\gamma > 0$, not for $\gamma \leq 0$

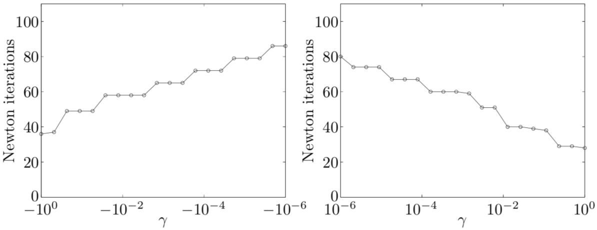


Example - Phase I Method

linear feasibility problems:

$$Ax \leq b + \gamma \Delta b$$

- $A \in \mathbb{R}^{50 \times 20}$
- strictly feasible for $\gamma > 0$, not for $\gamma \leq 0$

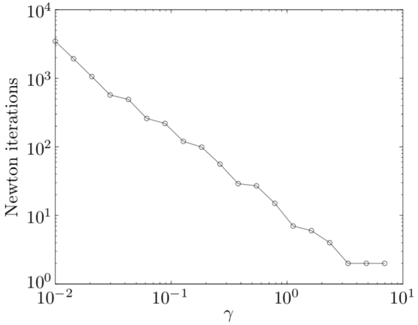


number of iterations roughly proportional to $\log(1/|\gamma|)$

Example - Infeasible Start Newton Method

infeasible start Newton method for the following problem:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} \log s_i \\ \text{subject to} & Ax + s = b + \gamma \Delta b \end{array}$$



# of iterations roughly proportional to $1/|\gamma|$ (worse than Phase I)

Outline

Problems With Generalized Inequalities

problems with generalized inequalities:

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

- same assumptions as before
- examples of great interests
  - K_i is second-order cone: SOCP
  - K_i is positive semidefinite cone: SDP

Generalized Logarithms

(standard) logarithm for nonnegative orthant $K = \mathbb{R}^n_+$:

$$\psi(x) = \sum_{i=1}^{n} \log x_i$$

generalized logarithm for positive semidefinite cone $K = \mathbb{S}^n_+$:

$$\psi(X) = \log \det X$$

for second-order cone $K = \left\{ x \in \mathbb{R}^{n+1} \mid \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \leq x_{n+1} \right\}$:

$$\psi(x) = \log \left(x_{n+1}^2 - \sum_{i=1}^{n} x_i^2 \right)$$

Properties of Generalized Logarithms

general properties:

$$\nabla^2\psi(x) \prec 0, \quad \nabla\psi(x) \succeq_{K^*} 0, \quad x^T\nabla\psi(x) = \theta$$

nonnegative orthant $K = \mathbb{R}^n_+$: $\psi(x) = \sum_{i=1}^n \log x_i$

$$\nabla\psi(x) = (1/x_1, \ldots, 1/x_n), \quad x^T\nabla\psi(x) = \theta$$

positive semidefinite cone $K = \mathbb{S}^n_+$: $\psi(X) = \log \det X$

$$\nabla\psi(X) = X^{-1}, \quad \mathbf{tr}(X\nabla\psi(X)) = n$$

second-order cone $K = \left\{x \in \mathbb{R}^{n+1} \mid \left(\sum_{i=1}^n x_i^2\right)^{1/2} \leq x_{n+1}\right\}$:

$$\nabla\psi(x) = \frac{2}{x_{n+1}^2 - \sum_{i=1}^n x_i^2} \begin{bmatrix} -x_1 \\ \cdots \\ -x_n \\ x_{n+1} \end{bmatrix}, \quad x^T\nabla\psi(x) = 2$$

Logarithmic Barrier and Central Path

logarithmic barrier for $f_i(x) \preceq_{K_i} 0, \; i = 1, \ldots, m$:

$$\phi(x) = -\sum_{i=1}^{m} \psi_i(-f_i(x)), \quad \mathbf{dom}\phi = \{x \mid f_i(x) \prec_{K_i} 0, \; i = 1, \ldots, m\}$$

- ψ_i is generalized logarithm for K_i
- ϕ convex, twice continuously differentiable

central path: $\{x^\star(t) \mid t > 0\}$ where $x^\star(t)$ is the solution to

$$\begin{array}{ll} \text{minimize} & tf_0(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

Dual Points on Central Path

$x^\star(t)$ satisfies

$$t\nabla f_0(x) + \sum_{i=1}^{m} Df_i(x)^T \nabla\psi(-f_i(x)) + A^T\hat{\nu} = 0$$

dual variables:

$$\lambda^\star(t) = \frac{1}{t}\nabla\phi_i(-f_i(x^\star(t))), \quad \nu^\star(t) = \frac{\hat{\nu}}{t}$$

duality gap:

$$f_0(x^\star(t)) - g(\lambda^\star(t), \nu^\star(t)) = \frac{\sum_{i=1}^{m}\theta_i}{t}$$

Barrier Method For Generalized Inequalities

the barrier method with strictly feasible starting point:

- strictly feasible starting point x, $t > 0$, $\mu > 1$, tolerance $\epsilon > 0$
- repeat the following steps
  1. centering step: find $x^\star(t)$ by minimizing $tf_0 + \phi$ s.t. $Ax = b$
  2. update $x := x^\star(t)$
  3. quit if $(\sum_{i=1}^m \theta_i)/t < \epsilon$
  4. increase $t := \mu t$

features:

- only difference is duality gap $(\sum_{i=1}^m \theta_i)/t$ (instead of m/t)
- number of outer iterations

$$1 + \left\lceil \frac{\log \left(\sum_{i=1}^m \theta_i\right)/(\epsilon t^{(0)})}{\log \mu} \right\rceil$$

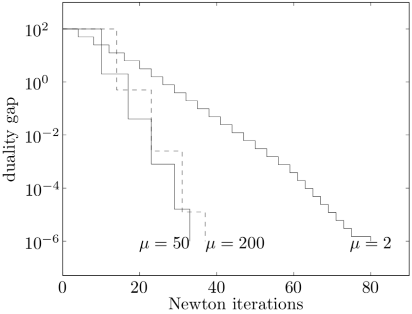
- convergence analysis similar
- same phase I method for finding strictly feasible points

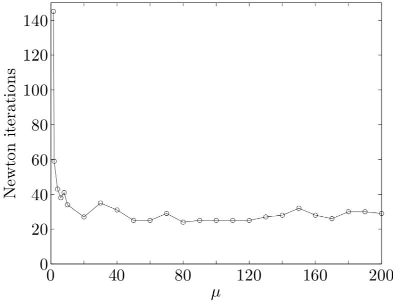
Examples – SOCP

SOCP:

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \le c_i^T x + d_i, \quad i = 1, \ldots, m \end{array}$$

with $m = 50$, $x \in \mathbb{R}^{50}$, $A_i \in \mathbb{R}^{5\times 50}$



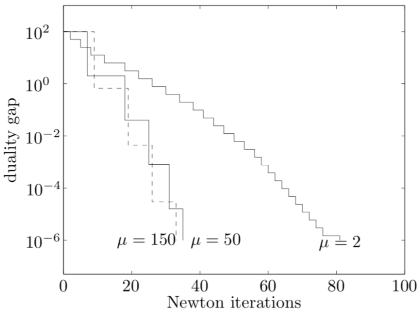


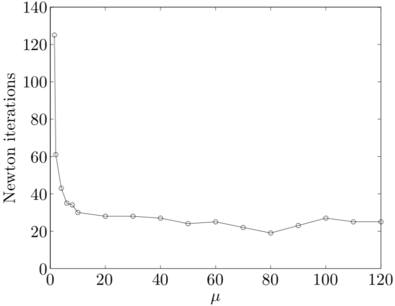
Examples – SDP

SDP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots x_n F_n + G \preceq 0 \end{array}$$

with $x \in \mathbb{R}^{100}$, $F_i, G \in \mathbb{S}^{100}$



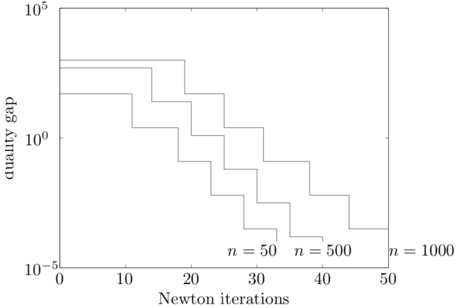


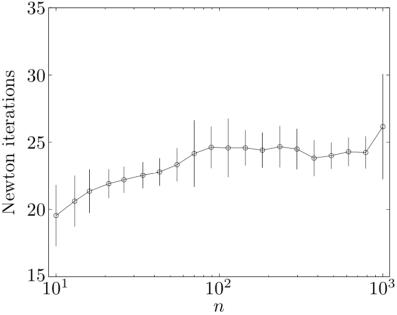
Examples – Scalability

a special SDP:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T x \\ \text{subject to} & A + \mathbf{diag}(x) \succeq 0 \end{array}$$

with $x \in \mathbb{R}^n$, $A \in \mathbb{S}^n$





Outline

Primal-Dual Interior-Point Methods

advantages over barrier methods:

- more efficient – no distinction between outer / inner iterations
- primal and dual variables updated at each iteration
- can start at infeasible points (for equality constraints)
- converge faster (empirically observed)

Primal-Dual Search Directions

recall the modified KKT conditions:

$$\begin{aligned} \nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + A^T \nu &= 0 \\ -\lambda_i f_i(x) &= 1/t, \; i = 1, \ldots, m \\ Ax &= b \end{aligned}$$

$n + m + p$ equations in variables $(x, \lambda, \nu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$

Primal-Dual Search Directions

residual $r_t(x, \lambda, \nu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$:

$$r_t(x, \lambda, \nu) = \begin{bmatrix} \nabla f_0(x) + Df(x)^T \lambda + A^T \nu \\ -\mathbf{diag}(\lambda) f(x) - (1/t)\mathbf{1}_m \\ Ax - b \end{bmatrix} \triangleq \begin{bmatrix} r_{\text{dual}} \\ r_{\text{cent}} \\ r_{\text{pri}} \end{bmatrix}$$

where we have

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \quad Df(x) = \begin{bmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{bmatrix}$$

$(x^\star(t), \lambda^\star(t), \nu^\star(t))$ satisfy

$$r_t(x^\star(t), \lambda^\star(t), \nu^\star(t))) = 0, \quad f_i(x^\star(t)) < 0, \ i = 1, \ldots, m$$

with duality gap m/t

Primal-Dual Search Directions

solve $r_t(x, \lambda, \nu) = 0$ through first-order Taylor approximation

given $y = (x, \lambda, \nu)$, find $\Delta y = (\Delta x, \Delta\lambda, \Delta\nu)$ such that

$$r_t(y + \Delta y) \approx r_t(y) + Dr_t(y)\Delta y = 0$$

specifically, we have

$$\begin{bmatrix} \nabla^2 f_0(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T \\ -\mathbf{diag}(\lambda) Df(x) & -\mathbf{diag}(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{bmatrix} = - \begin{bmatrix} r_{\mathrm{dual}} \\ r_{\mathrm{cent}} \\ r_{\mathrm{pri}} \end{bmatrix}$$

primal-dual search directions Δy_{pd}: solution to the above

main difference from barrier method: update $\Delta\lambda$

Primal-Dual Interior-Point Method

surrogate duality gap

$$\hat{\eta}(x, \lambda) = -f(x)^T \lambda$$

equal to true duality gap when x and λ are primal and dual feasible

primal-dual interior-point method:

- given x with $f(x) < 0$, $\lambda > 0$, $\mu > 1$, tolerance $\epsilon_{\text{feas}}, \epsilon > 0$
- repeat the following steps
  1. set $t := \mu m / \hat{\eta}$
  2. compute primal-dual search direction Δy_{pd}
  3. backtracking line search on λ, $f(x)$, and $\|r_t\|_2$
     1. start with $s := 0.99 \cdot \sup\{s \in [0,1] \mid \lambda + s\Delta\lambda \geq 0\}$
     2. continue $s := \beta s$ until $f(x + s\Delta x) < 0$
     3. continue $s := \beta s$ until $\|r_t(y + s\Delta y_{\text{pd}}\|_2 > (1 - \alpha s)\|r_t(y)\|_2$,
  4. update $y := y + s\Delta y_{\text{pd}}$
- until $\|r_{\text{pri}}\|_2 \leq \epsilon_{\text{feas}}$, $\|r_{\text{dual}}\|_2 \leq \epsilon_{\text{feas}}$, and $\hat{\eta} \leq \epsilon$

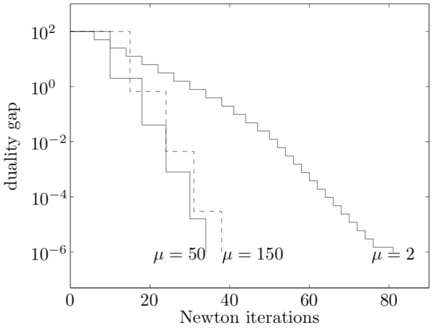
Examples – LP in Inequality Form

LP in inequality form

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

with $A \in \mathbb{R}^{100 \times 50}$

barrier method:



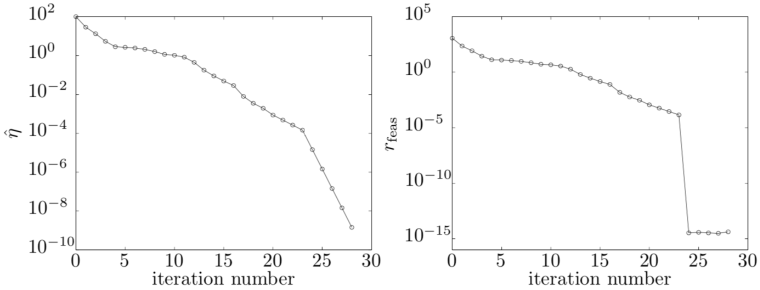
Examples – LP in Inequality Form

LP in inequality form

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

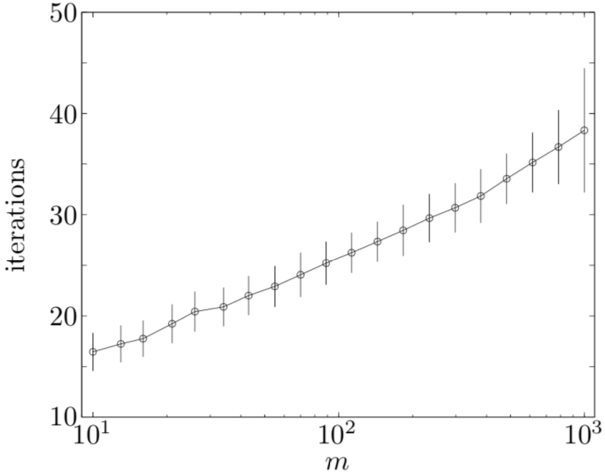
with $A \in \mathbb{R}^{100 \times 50}$

primal-dual interior-point method:



Examples – Scalability

for the LP, fix $n = 2m$ and let m increase



Outline

Implementation Issues

main effort in barrier method:

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ \nu_{\text{nt}} \end{bmatrix} = - \begin{bmatrix} g \\ 0 \end{bmatrix}$$

where

$$H = t\nabla^2 f_0(x) + \sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

and

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

solve a linear system of size $(n+p)$

complexity $O((n+p)^3)$ in general

Exploit Structures to Reduce Complexity

sparse problem:

- objective and constraint functions depend on a few variables
- H is likely to be sparse if m small
- A is sparse

use the structure to reduce complexity

customize the method for independent problems

Examples – Standard Form LP

LP in standard form:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

centering problem:

$$\begin{array}{ll} \text{minimize} & tc^T x - \sum_{i=1}^{n} \log x_i \\ \text{subject to} & Ax = b \end{array}$$

Newton steps:

$$\begin{bmatrix} \mathbf{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ \nu_{\text{nt}} \end{bmatrix} = - \begin{bmatrix} -tc + \mathbf{diag}(x)^{-1}\mathbf{1} \\ 0 \end{bmatrix}$$

Examples – Standard Form LP

solving Newton steps:

$$\begin{bmatrix} \mathbf{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ \nu_{\text{nt}} \end{bmatrix} = \begin{bmatrix} -tc + \mathbf{diag}(x)^{-1}\mathbf{1} \\ 0 \end{bmatrix}$$

determine Δx_{nt} analytically from ν_{nt}:

$$\begin{aligned} \Delta x_{\text{nt}} &= \mathbf{diag}(x)^2 \left(-tc + \mathbf{diag}(x)^{-1}\mathbf{1} - A^T \nu_{\text{nt}} \right) \\ &= -t\mathbf{diag}(x)^2 c + x - \mathbf{diag}(x)^2 A^T \nu_{\text{nt}} \end{aligned}$$

solve ν_{nt}:

$$A\,\mathbf{diag}(x)^2 A^T \nu_{\text{nt}} = -tA\mathbf{diag}(x)^2 c + b$$

Examples – ℓ_1-Norm Approximation

ℓ_1-norm approximation problem:

$$\text{minimize} \quad \|Ax - b\|_1$$

with $A \in \mathbb{R}^{m \times n}$

equivalent LP:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T y \\ \text{subject to} & \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{array}$$

with optimization variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$

Examples – ℓ_1-Norm Approximation

Newton steps:

$$\begin{bmatrix} A^T(D_1+D_2)A & -A^T(D_1-D_2) \\ -(D_1-D_2)A & D_1+D_2 \end{bmatrix} \begin{bmatrix} \Delta x_{\rm nt} \\ \Delta y_{\rm nt} \end{bmatrix} = - \begin{bmatrix} A^T g_1 \\ g_2 \end{bmatrix}$$

where D_1 and D_2 are diagonal matrices (expressions omitted)

by eliminating $\Delta y_{\rm nt}$, we have

$$A^T D A \Delta x_{\rm nt} = -A^T g$$

with

$$D = 4 D_1 D_2 (D_1 + D_2)^{-1}$$

and

$$g = g_1 + (D_1 - D_2)(D_1 + D_2)^{-1} g_2$$

then get $\Delta y_{\rm nt}$ through

$$\Delta y_{\rm nt} = (D_1 + D_2)^{-1}(-g_2 + (D_1 - D_2) A \Delta x_{\rm nt})$$

Examples – Network Rate Optimization

network rate optimization problem:

- n flows (e.g., traffic, commodity)
- L links with capacities

optimization problem:

$$\begin{array}{ll} \text{maximize} & U(x) = U_1(x_1) + \cdots + U_n(x_n) \\ \text{subject to} & Ax \le c, \;\; x \ge 0 \end{array}$$

with $A \in \{0,1\}^{m \times n}$ is the incident matrix

$$A_{ij} = \begin{cases} 1 & \text{flow } j \text{ pass through link } i \\ 0 & \text{otherwise} \end{cases}$$

Examples – Network Rate Optimization

centering problem:

$$\text{minimize} \quad -tU(x) - \sum_{i=1}^{L} \log\left(c - Ax\right)_i - \sum_{j=1}^{n} \log x_j$$

Newton steps:

$$\left(D_0 + A^T D_1 A + D_2\right) \Delta x_{\text{nt}} = -g$$

where D_0, D_1 and D_2 are diagonal matrices (expressions omitted)

$$\left(D_0 + A^T D_1 A + D_2\right)_{ij} \neq 0$$

if and only if flows i and j share a link