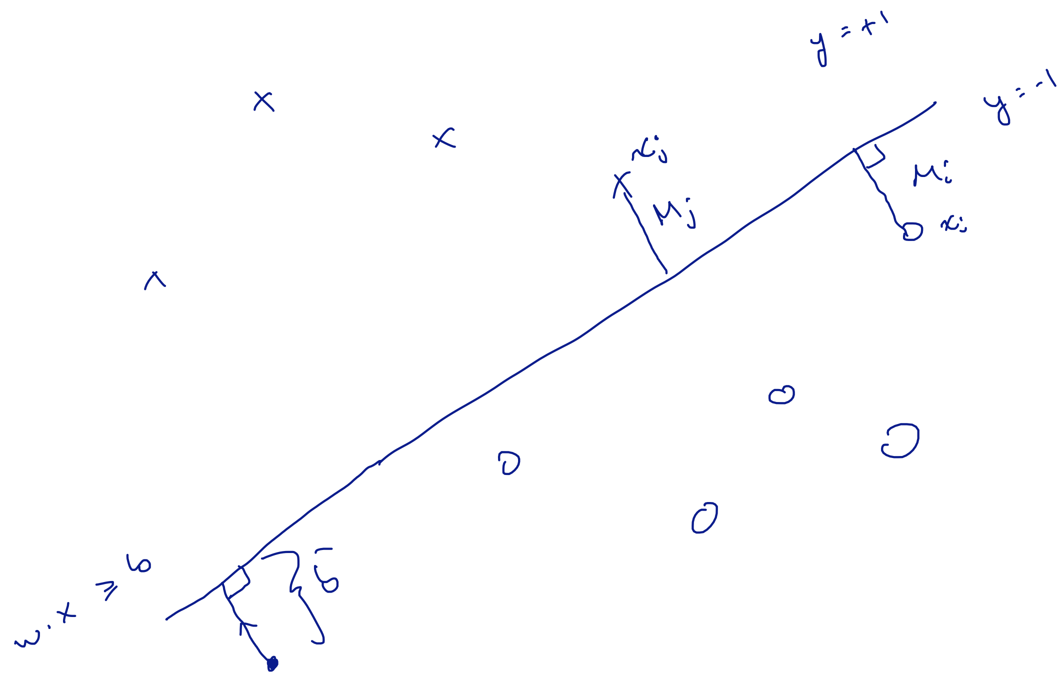
Interior-point methods

10-725 Optimization
Geoff Gordon
Ryan Tibshirani

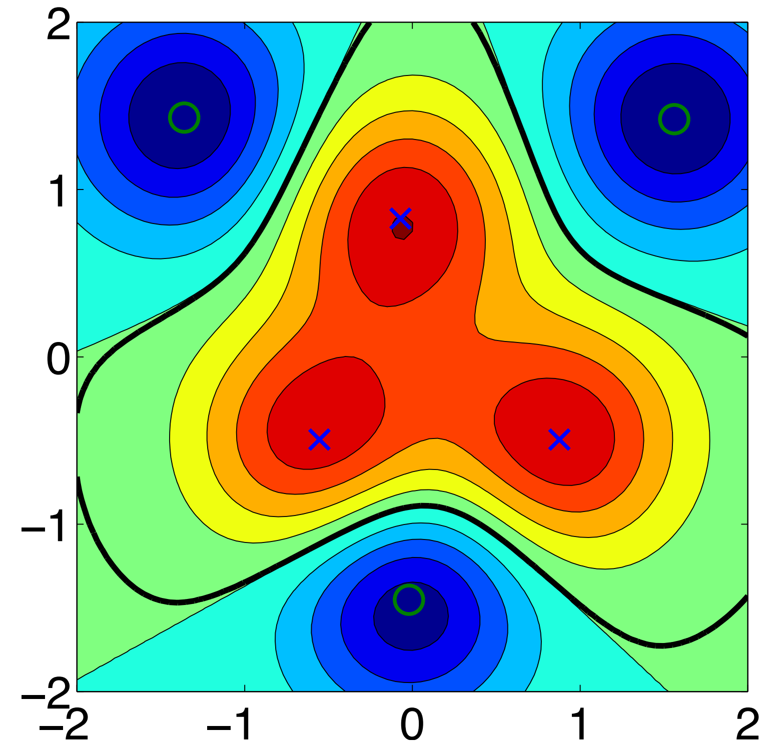
Review

- SVM duality
  - min $v^Tv/2 + 1^Ts$ s.t. $Av - yd + s - 1 \geq 0$ $s \geq 0$
  - max $1^T\alpha - \alpha^TK\alpha/2$ s.t. $y^T\alpha = 0$ $0 \leq \alpha \leq 1$
  - Gram matrix K
- Interpretation
  - support vectors & complementarity
  - reconstruct primal solution from dual



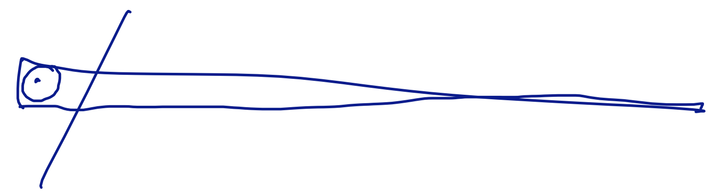
Review

- Kernel trick
  - high-dim feature spaces, fast
  - positive definite function
- Examples
  - polynomial
  - homogeneous polynomial
  - linear
  - Gaussian RBF



Review: LF problem $Ax + b \geq 0$

- Ball center
  - bad summary of LF problem

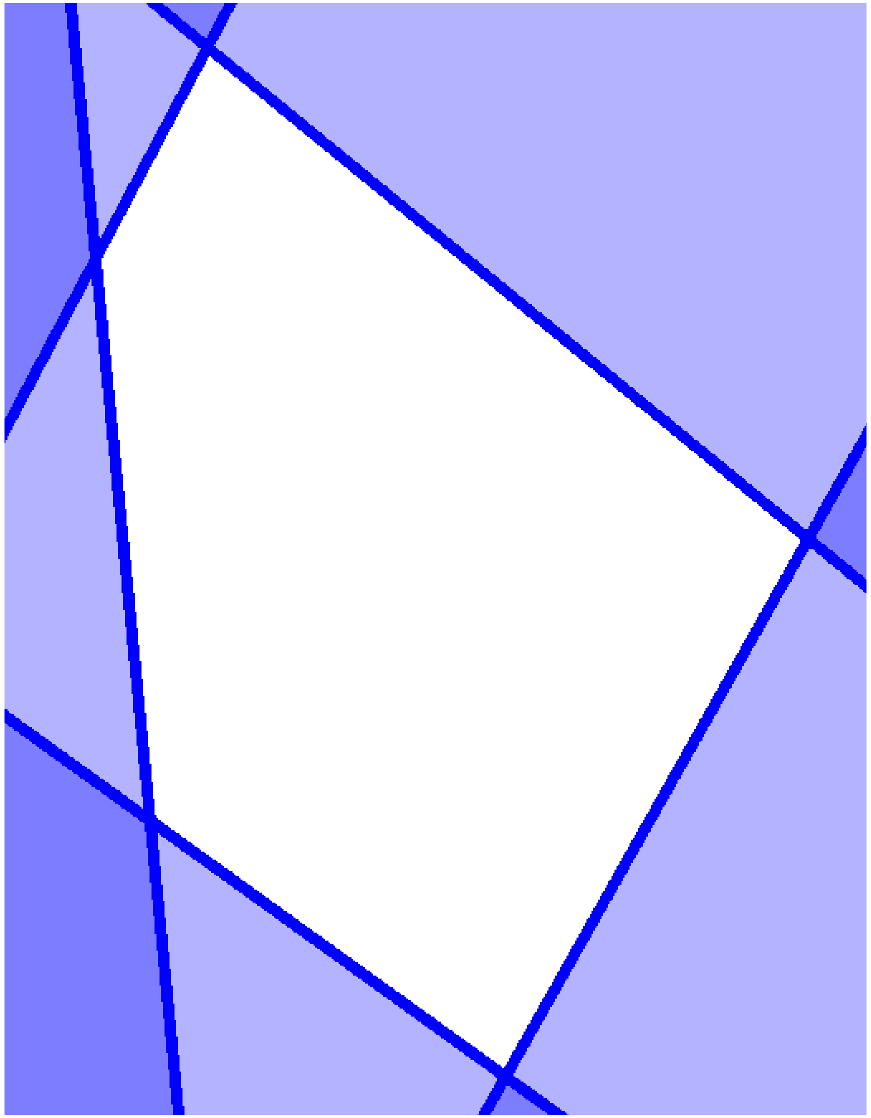


- Max-volume ellipsoid / ellipsoid center
  - good summary (1/n of volume), but expensive
- Analytic center of LF problem
  - maximize product of distances to constraints
  - $\min -\sum \ln(a_i^T x + b_i)$
- Dikin ellipsoid @ analytic center: not quite as good (just 1/m < 1/n), but much cheaper

Force-field interpretation

of analytic center

- Pretend constraints are repelling a particle
  - normal force for each constraint
  - force $\propto$ 1/distance
- Analytic center = equilibrium = where forces balance

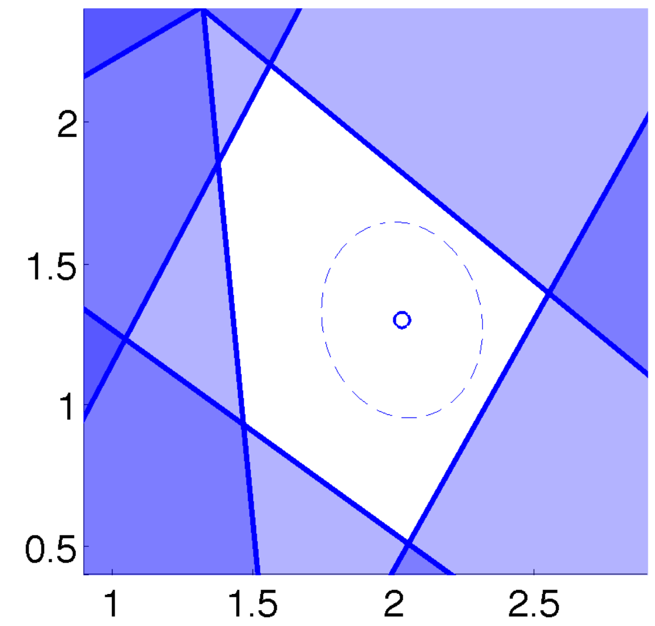


Newton for analytic center

- $f(x) = -\sum \ln(a_i^T x + b_i)$
  - df/dx =
  - d^2f/df^2 =

Dikin ellipsoid

- $E(x_0) = \{ x \mid (x - x_0)^T H (x - x_0) \leq 1 \}$
  - H = Hessian of log barrier at x_0
  - unit ball of Hessian norm at x_0
- $E(x_0) \subseteq X$ for any strictly feasible x_0
  - affine constraints can be just feasible
  - $E(x_0)$: as above, but intersected w/ affine constraints
- $\mathrm{vol}(E(x_{ac})) \geq \mathrm{vol}(X)/m$
  - weaker than ellipsoid center, but still very useful



$E(x_0) \subseteq X$

- $E(x_0) = \{ x \mid (x-x_0)^T H (x-x_0) \leq 1 \}$
  - $H = A^T S^{-2} A$
  - $S = \mathrm{diag}(s) = \mathrm{diag}(Ax_0 + b)$

$mE(x_0) \supseteq X$

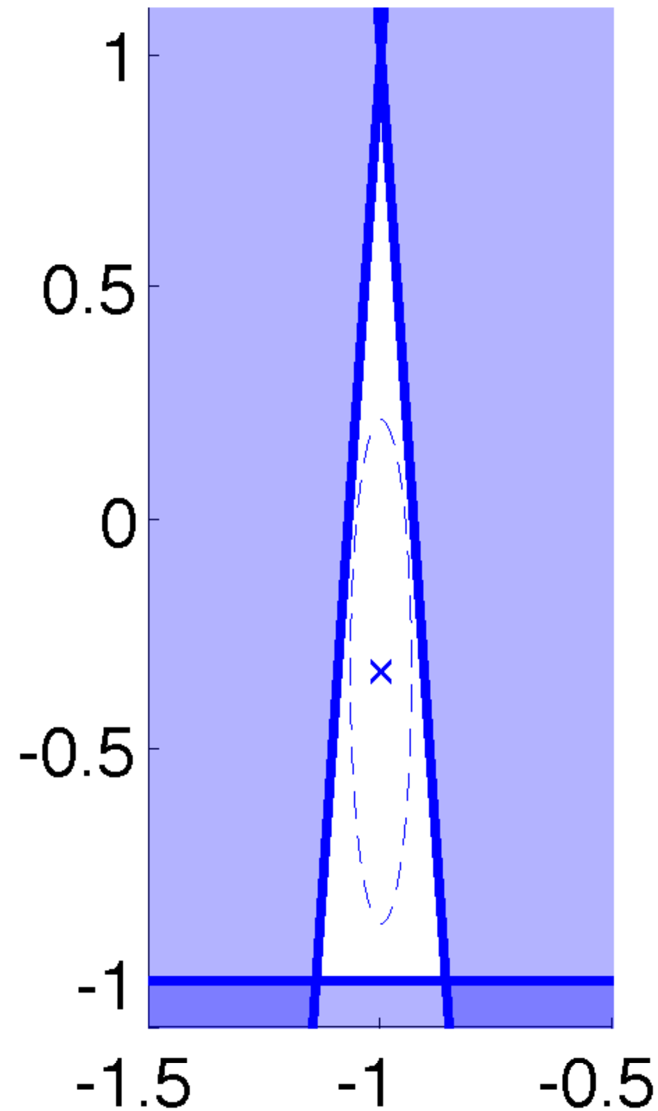
- Feasible point x: $Ax + b \geq 0$
- Analytic center x_{ac}: $A^T y = 0$ $y = 1./(Ax_{ac}+b)$
- Let $Y = \text{diag}(y_{ac})$, $H = A^T Y^2 A$; show:
  - $(x-x_{ac})^T H (x-x_{ac}) \leq m^2$ [+ m]

Combinatorics v. analysis

- Two ways to find a feasible point of $Ax+b \geq 0$
  - find analytic center—minimize a smooth function
  - find a feasible basis—combinatorial search

Bad conditioning? No problem.

- Analytic center & Dikin ellipsoids invariant to affine xforms $w = Mx+q$
  - $W = \{ w \mid AM^{-1}(w-q) + b \geq 0 \}$
- Can always xform so that a ball takes up ≥ vol(Y)/m
  - Dikin ellipsoid @ac → sphere



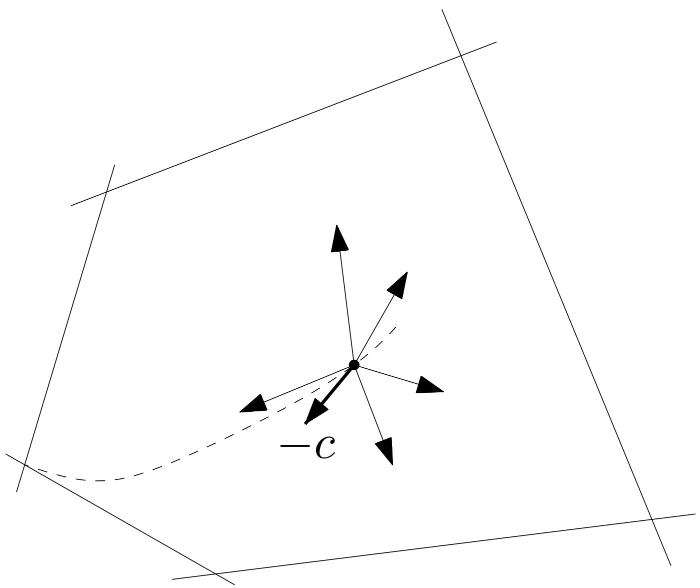
LF→LP: the central path

- Analytic center was for: find x st $Ax + b \geq 0$
- Now: min $c^T x$ st $Ax + b \geq 0$
- Same trick:
  - min $f_t(x) = c^T x - (1/t) \sum \ln(a_i^T x + b_i)$
    - parameter $t > 0$
  - central path =
    - $t \to 0$: $t \to \infty$:

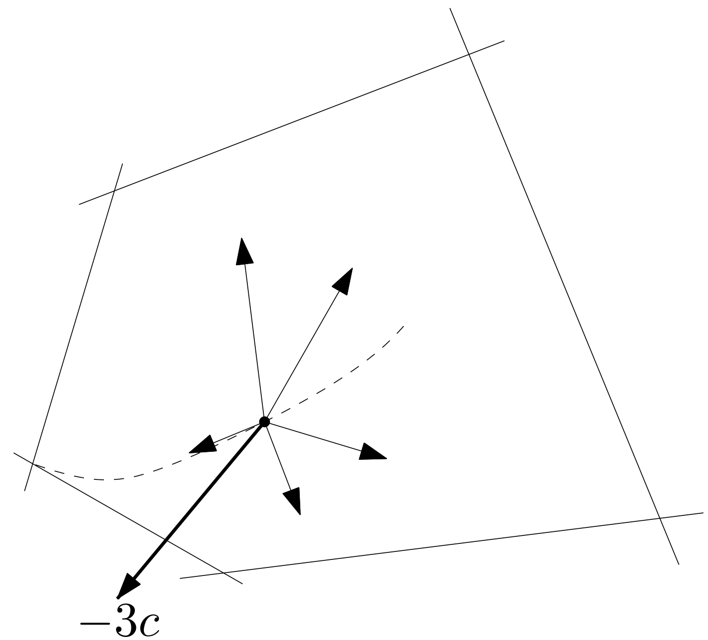
Force-field interpretation

of central path

- Force along objective; normal forces for each constraint



t=1

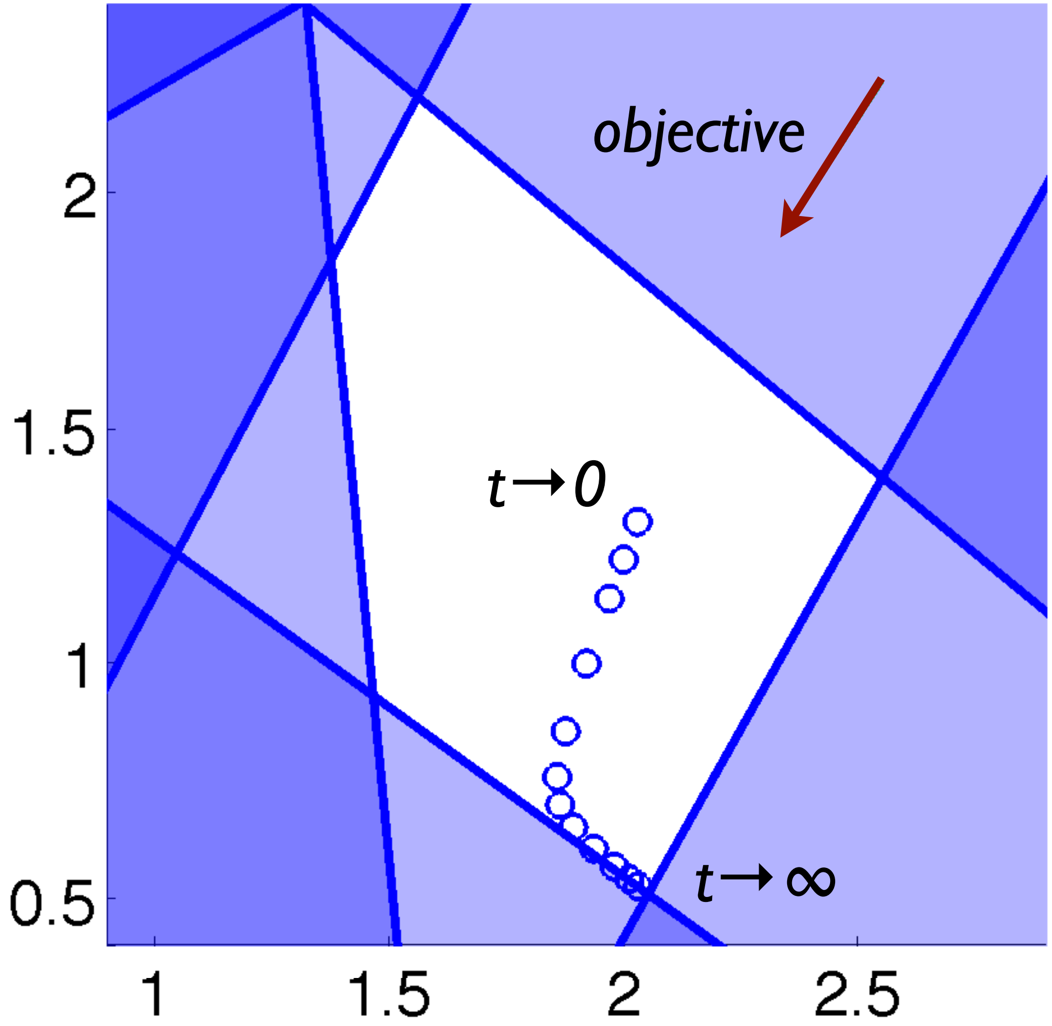


t=3

Newton for central path

- min $f_t(x) = c^T x - (1/t) \sum \ln(a_i^T x + b_i)$
  - df/dx =
  - d^2f/dx^2 =

Central path example



New LP algorithm?

- Set $t=10^{12}$. Find corresponding point on central path by Newton's method.
  - worked for example on previous slide!
  - but has convergence problems in general

- Alternatives?

Constraint form of central path

- $\min -\sum \ln s_i$ st $Ax + b \geq 0 \quad c^T x \leq \lambda$
- $\exists$ a 1-1 mapping $\lambda(t)$ w/ $x(\lambda(t)) = x(t)\ \forall t>0$
  - but this form is slightly less convenient since we don't know minimal feasible value of λ or maximal nontrivial value of λ

Dual of central path

- min $c^T x - (1/t) \sum \ln s_i$ st $Ax + b = s \geq 0$
    - $\min_{x,s} \max_y L(x,s,y) = c^T x - (1/t) \sum \ln s_i + y^T(s - Ax - b)$

Primal-dual correspondence

- Primal and dual for central path:
  - min $c^T x - (1/t) \sum \ln s_i$ st $Ax + b = s \geq 0$
  - max $(m \ln t)/t + m/t + (1/t) \sum \ln y_i - y^T b$ st $A^T y = c \quad y \geq 0$
- $L(x,s,y) = c^T x - (1/t) \sum \ln s_i + y^T(s - Ax - b)$
  - grad wrt s:
  - to get x:

Duality gap

- At optimum:
  - primal value $c^T x - (1/t) \sum \ln s_i$ =
    dual value $(m \ln t)/t + m/t + (1/t) \sum \ln y_i - y^T b$
  - $s \circ y = te$

Primal-dual constraint form

- Primal-dual pair:
  - min $c^T x$ st $Ax + b \geq 0$
  - max $-b^T y$ st $A^T y = c \quad y \geq 0$
- KKT:
  - $Ax + b \geq 0$ (primal feasibility)
  - $y \geq 0 \quad A^T y = c$ (dual feasibility)
  - $c^T x + b^T y \leq 0$ (strong duality)
  - …or, $c^T x + b^T y \leq \lambda$ (relaxed strong duality)

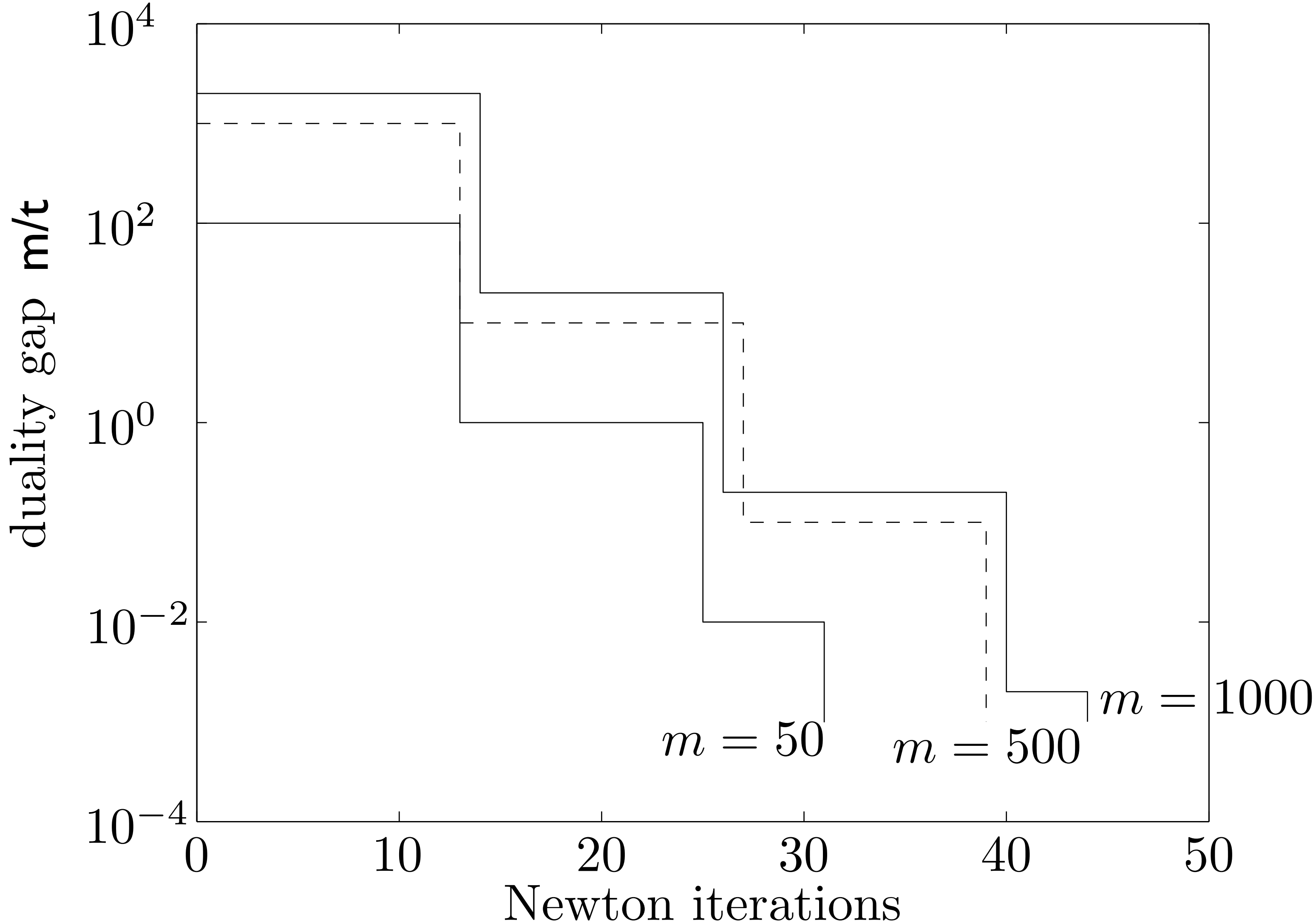
Analytic center of relaxed KKT

- Relaxed KKT conditions:
  - $Ax + b = s \geq 0$
  - $y \geq 0$
  - $A^T y = c$
  - $c^T x + b^T y \leq \lambda$
- Central path = {analytic centers of relaxed KKT}

Algorithm

- $t := 1, y := 1^m, x := 0^n$ $[s := 1^m]$
- Repeat
  - Use infeasible-start Newton to find point on dual central path
  - Recover primal (s,x); gap $c^T x + b^T y = m/t$
    - $s = 1./ty$ $x = A\backslash(s-b)$ [have already (Newton)]
  - $t := \alpha t$ $(\alpha > 1)$

Example



duality gap m/t
Newton iterations
10^4
10^2
10^0
10^{-2}
10^{-4}
0
10
20
30
40
50
$m = 50$
$m = 500$
$m = 1000$