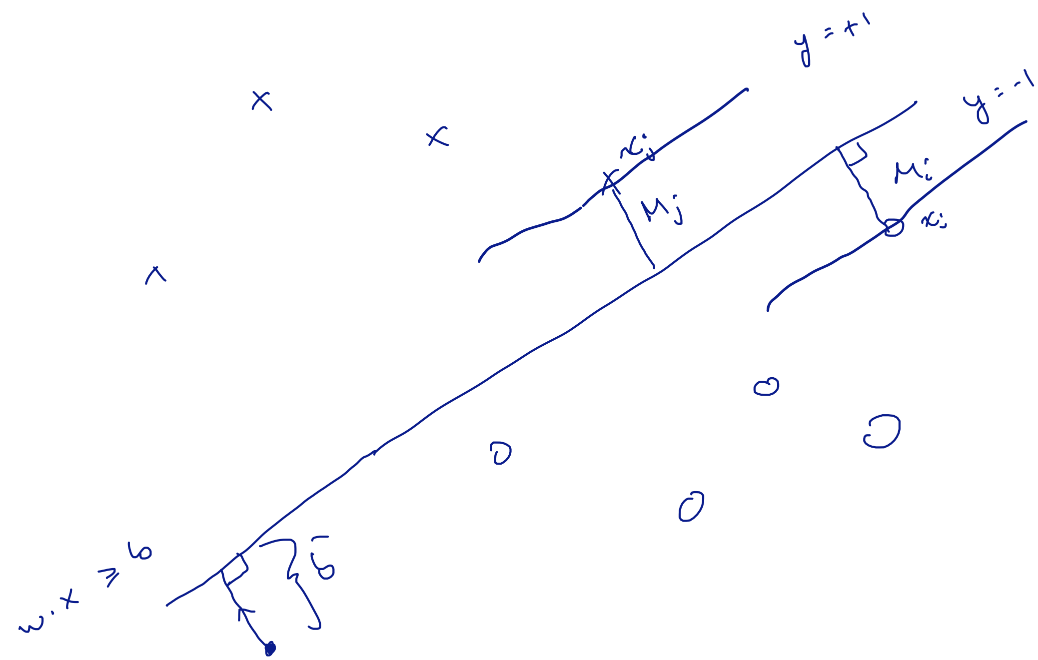
Interior-point methods

10-725 Optimization
Geoff Gordon
Ryan Tibshirani

Review

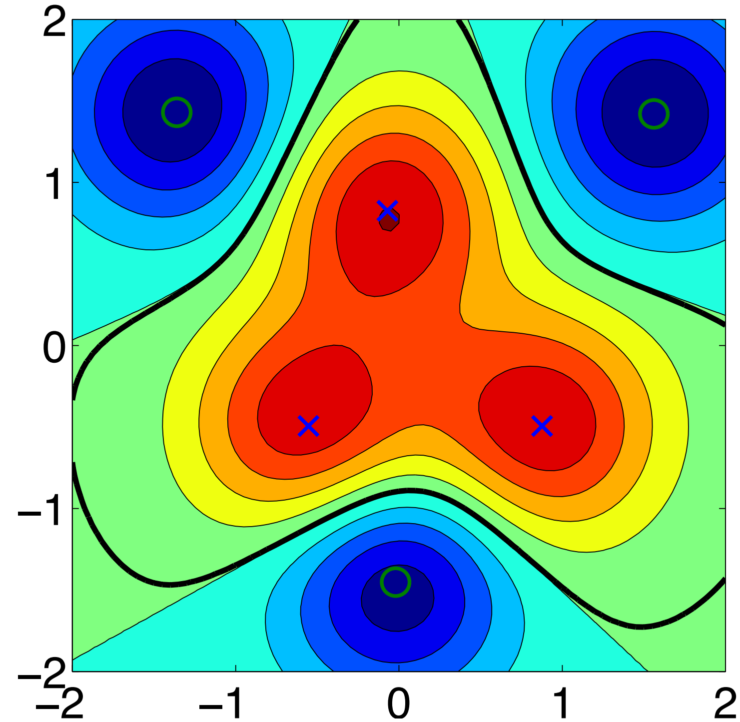
- SVM duality C
  - min $v^Tv/2 + 1^Ts$ s.t. $Av - yd + s - 1 \geq 0$ $s \geq 0$
  - max $1^T\alpha - \alpha^TK\alpha/2$ s.t. $y^T\alpha = 0$ $0 \leq \alpha \leq 1$
  - Gram matrix K
- Interpretation
  - support vectors & complementarity
  - reconstruct primal solution from dual



Review

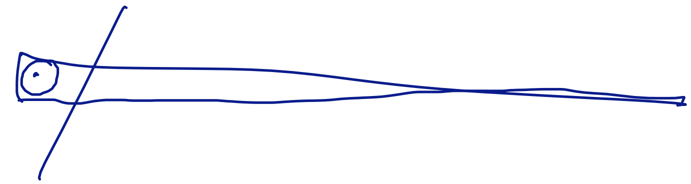
$k(u_i, u_j)$ $\phi(u_i) = x_i$

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

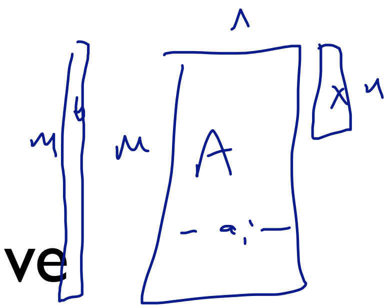


$a_i^T x + b_i \geq 0$

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



- Ball center
  - bad summary of LF problem



- Max-volume ellipsoid / ellipsoid center
  - good summary ($1/n^n$ of volume), but expensive
- Analytic center of LF problem
  - maximize product of distances to constraints

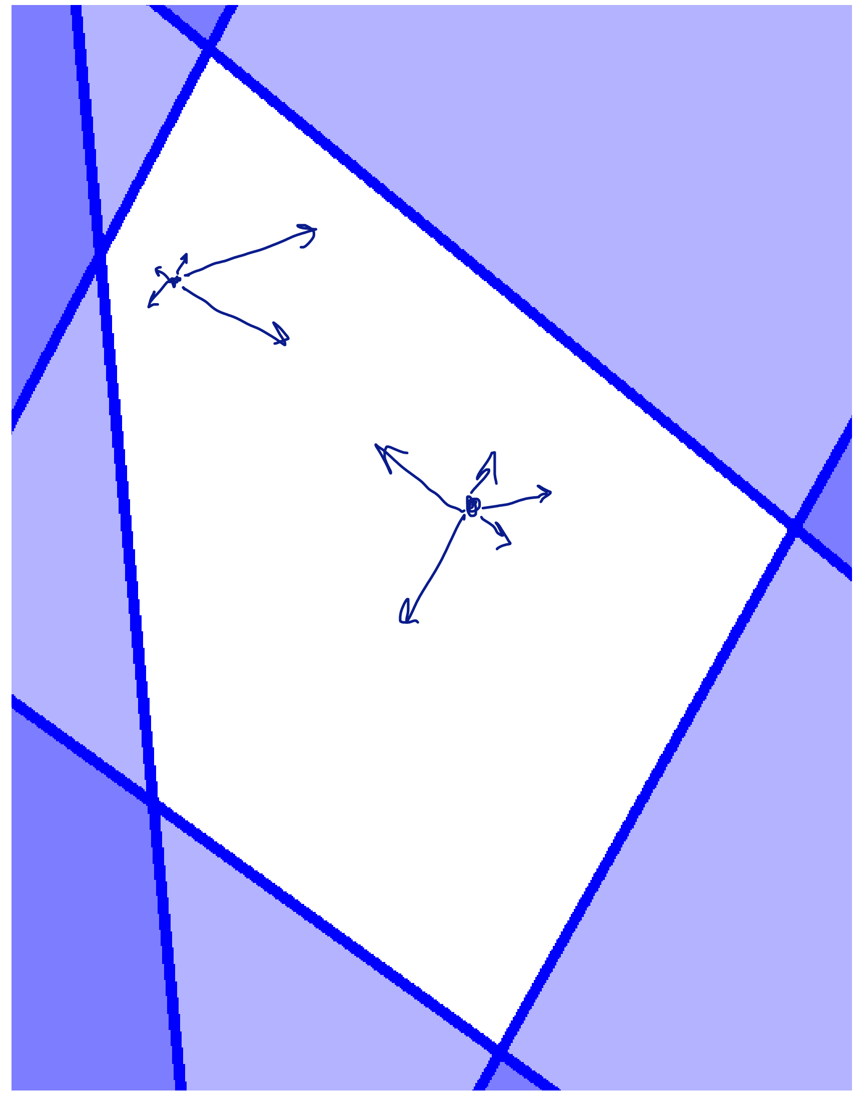
$(a_i \cdot x + b_i) / \|a_i\|$

  - min $-\sum \ln(a_i^T x + b_i)$
- Dikin ellipsoid @ analytic center: not quite as good (just $1/m^n < 1/n^n$), 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

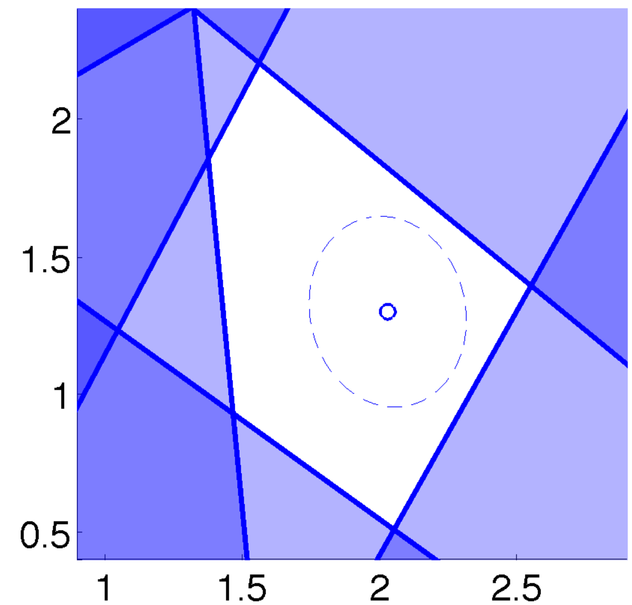
- $f(x) = -\sum \ln(a_i^T x + b_i)$ $s_i = a_i^T x + b_i$ $y_i = 1/s_i$ $S = \mathrm{diag}(s)$
  - df/dx = $\sum_i -\left(1/a_i^T x + b_i\right) a_i$ $-A^T y = 0$
  - $d^2f/df^2 =$ $+\sum_i (a_i^T x + b_i)^{-2} a_i a_i^T$ $H = A^T S^{-2} A = A^T Y^2 A$

Newton $\Delta x = (A^T S^{-2} A)^{-1} A^T y$

Dikin ellipsoid

$\|x\|_H = \sqrt{x^T H x}$

- $E(x_0) = \{ x \mid (x-x_0)^T H (x-x_0) \le 1 \}$
  - 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
- $\mathrm{vol}(E(x_{ac})) \ge \mathrm{vol}(X)/m^n$
  - weaker than ellipsoid center, but still very useful



$E(x_0) \subseteq X$

$s_i = a_i^T x_0 + b_i$

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

$H = \sum_i s_i^{-2} a_i a_i^T$

$\sum_i (a_i^T(x-x_0)/s_i)^2 \leq 1$

$-1 \leq a_i^T(x-x_0)/s_i \leq 1 \quad \forall i$

$-s_i \leq a_i^T(x-x_0)$

$0 \leq a_i^T x + b_i$

$mE(\overset{x_{ac}}{\cancel{x_0}}) \supseteq X$

$y_i = 1/a_i^T x_{ac} + b_i$

- 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]

$$\sum_i y_i^2 (a_i^T (x - x_{ac}))^2 = \sum_i y_i^2 (a_i^T x + b_i - 1/y_i)^2$$

$$= \sum_i y_i^2 \left((a_i^T x + b_i)^2 + \frac{1}{y_i^2} - 2(a_i^T x + b_i)/y_i \right)$$

$$\leq m + \sum y_i^2 (a_i^T x + b_i)^2 \leq m + \left(\sum_i y_i (a_i^T x + b_i)\right)^2 \overset{?}{\leq} m + m^2$$

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.

M^{-1} exists

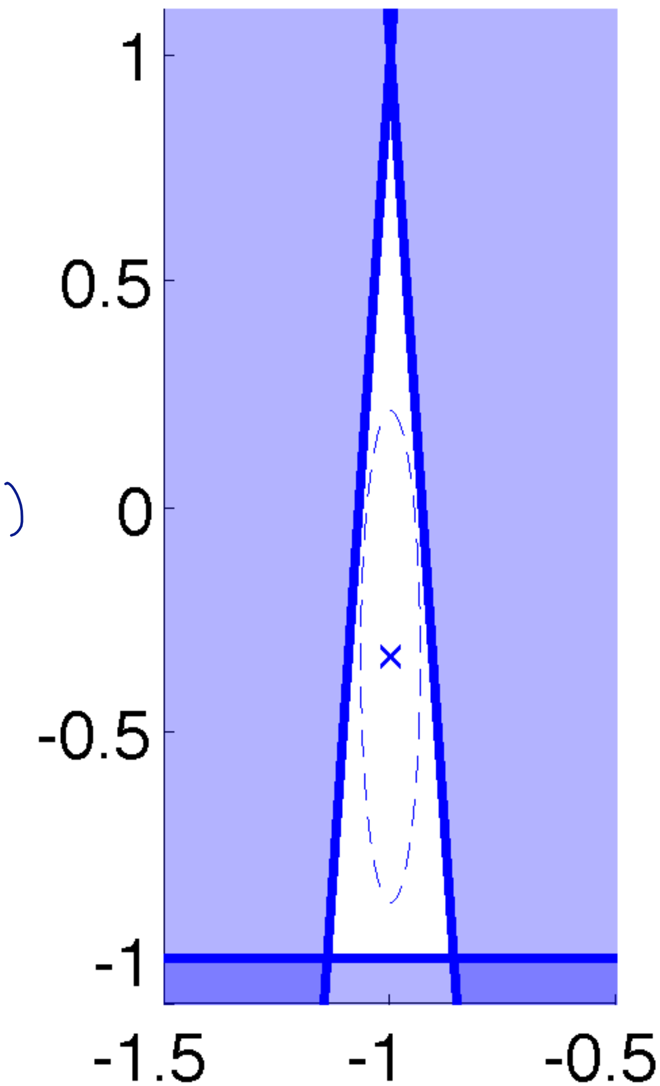
- Analytic center & Dikin ellipsoids invariant to affine xforms w = Mx+q $H_x = R^T R$
  - $W = \{ w \mid AM^{-1}(w-q) + b \geq 0 \}$

$w_{ac} = M x_{ac} + q$ $H_w = M^{-T} H_x M^{-1}$

$(M = R$ $H_w = R^{-T} R^T R R^{-1})$

$= I$

- Can always xform so that a ball takes up ≥ vol(Y)/m
  - Dikin ellipsoid @ac → sphere



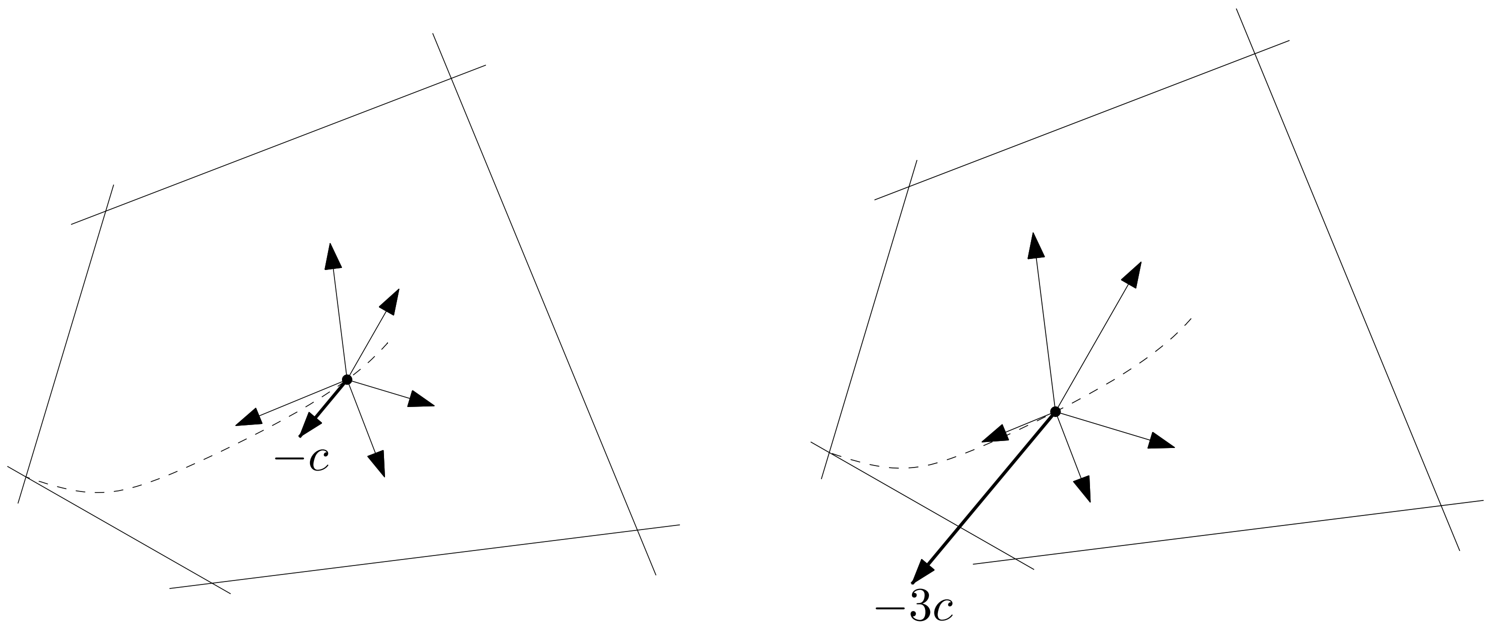
LF→LP: the central path

- Analytic center was for: find x st $Ax + b \geq 0$
- Now: min c^Tx st $Ax + b \geq 0$
- Same trick:
  - min $f_t(x) = c^Tx - (1/t) \sum \ln(a_i^Tx + b_i)$
    - parameter $t > 0$ $\quad \propto -H \backslash c$
  - central path = $\{x(t) \mid t > 0\}$
    - $t \to 0$: analytic center $\quad t \to \infty$: LP opt

Force-field interpretation

of central path

- Force along objective; normal forces for each constraint

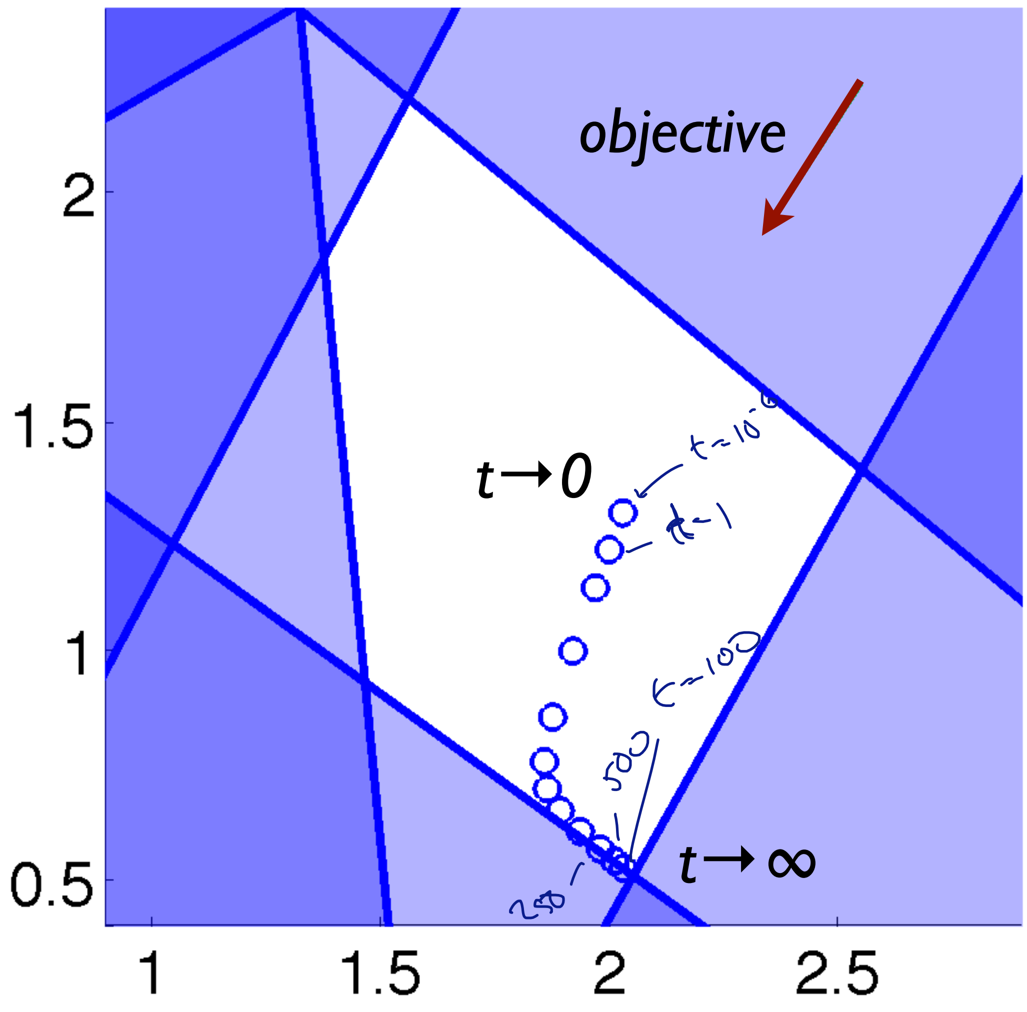


t=1 *t=3*

Newton for central path

- min $f_t(x) = c^T x - (1/t) \sum \ln(a_i^T x + b_i)$
  - $df/dx =$ $c - (1/t) \sum 1/s_i \, a_i$
    $\left(\sum s_i^{-2} a_i a_i^T\right) \Delta x = \sum \frac{a_i}{s_i} - tc$
  - $d^2f/dx^2 =$ $\frac{1}{t} \sum s_i^{-2} a_i a_i^T$

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

$\nabla_x :\ c - A^T y \qquad c = A^T y$

$\nabla_s :\ -1/ts + y \qquad s = 1/ty$

$\max_y \ \frac{m}{t} - y^T b + \frac{1}{t} \sum (\ln y_i + \ln t)$

$\max_y \left[\frac{m}{t} + \frac{m \ln t}{t}\right] - y^T b + \frac{1}{t} \sum \ln y_i$

st $A^T y = c \qquad (y \geq 0)$

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: $y - 1/ts \Rightarrow s_i y_i = 1/t \qquad s \circ y = 1/t$
  - to get x: $Ax + b = 1/ty$

Duality gap

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

$$\sum \ln s_i y_i = \sum -\ln t$$

$$\sum \ln s_i + \ln y_i = -m \ln t$$

$$c^T x - y^T b = m/t$$