Dikin's Method

Stephen Boyd (with help from AJ Friend)

EE364b, Stanford University

Outline

Dikin's original method for LP

Generalized Dikin's method

Dikin's method

- a simple proto interior-point method, originally for solving LPs
- invented by Dikin in 1967, but ignored/unknown for decades
- also called affine scaling method
- has very simple interpretation

Standard form LP

primal problem:

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

dual:

$$\begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & c + A^T \nu \succeq 0 \end{array}$$

optimality conditions:

- primal feasibility: $Ax = b$, $x \succeq 0$
- dual feasibility: $c + A^T \nu \succeq 0$
- zero gap/ complementarity: $x^T(c + A^T \nu) = c^T x + b^T \nu = 0$

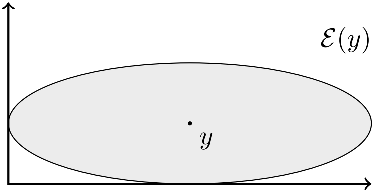
Dikin ellipsoid

- for $y \succ 0$, *Dikin ellipsoid* is

$$\mathcal{E}(y) = \{x \mid (x-y)^T H(x-y) \le 1\}, \qquad H = \mathbf{diag}(y)^{-2}$$

- $\mathcal{E}(y) \subset \mathbf{R}^n_+$; follows from

$$\sum_i \frac{(x_i - y_i)^2}{y_i^2} \le 1 \;\Rightarrow\; \frac{(x_i - y_i)^2}{y_i^2} \le 1 \;\Rightarrow\; |x_i - y_i| \le y_i \;\Rightarrow\; x_i \ge 0$$

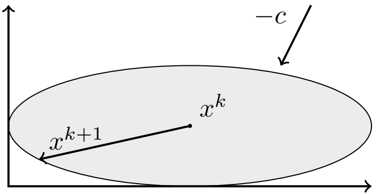


Dikin's method

- start with (strictly feasible) $x^0 \succ 0$, $Ax^0 = b$
- x^{k+1} is solution of

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{E}(x^k) \end{array}$$

- there's a simple formula for x^{k+1}
- maintains feasibility
- converges to solution



Dikin update

- with $H = \mathbf{diag}(x^k)^{-2}$, x^{k+1} is solution of

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & (x - x^k)^T H (x - x^k) \leq 1 \end{array}$$

- find $\Delta x^k = x^{k+1} - x^k$ via KKT system

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \mu\nu \end{bmatrix} = \begin{bmatrix} -\mu c \\ 0 \end{bmatrix},$$

  where μ is chosen after the solve to enforce the ellipsoid constraint
- more explicitly:

$$\begin{array}{rcl} \nu^k & = & -(AH^{-1}A^T)^{-1}AH^{-1}c \\ s^k & = & -H^{-1}(c + A^T\nu^k) \\ \mu^k & = & 1/\sqrt{s^{kT}Hs^k} \\ \Delta x^k & = & \mu^k s^k \end{array}$$

Comparison with barrier method

- barrier method 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 step Δx^k given by

$$\begin{bmatrix} H^k & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ t\nu \end{bmatrix} = \begin{bmatrix} -tc + \mathbf{diag}(x^k)^{-1}\mathbf{1} \\ 0 \end{bmatrix}$$

  where $H^k = \mathbf{diag}(x^k)^{-2}$

- with $t = \mu$, same as Dikin update except for centering term $\mathbf{diag}(x^k)^{-1}\mathbf{1}$

Stopping criteria

- Dikin iterates always primal feasible: $Ax^k = b$, $x^k \succeq 0$
- dual feasibility and zero duality gap only satisfied in the limit
- reasonable stopping criteria are

$$\min(c + A^T\nu^k) \geq -\epsilon_{\mathrm{df}}, \qquad |c^T x^k + b^T \nu^k| \leq \epsilon_{\mathrm{gap}}$$

  (second term is a *pseudo-gap*; it is only the true gap when $c + A^T\nu^k \succeq 0$)
- $\mu(c + A^T\nu^k) = -H^k \Delta x^k = -\left(\Delta x_1^k/(x_1^k)^2, \ldots, \Delta x_n^k/(x_n^k)^2\right)$
- stopping criteria can be written

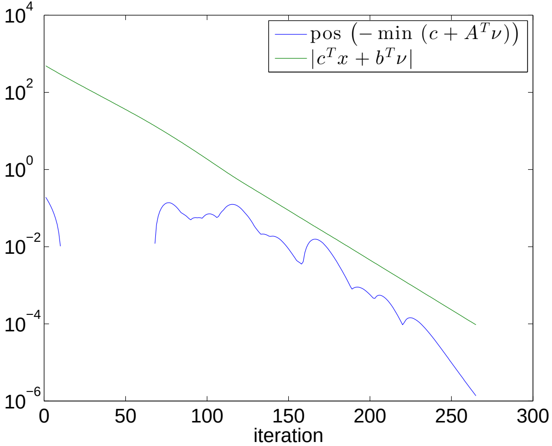
$$\max_i \frac{1}{\mu}\frac{\Delta x_i^k}{(x_i^k)^2} \leq \epsilon_{\mathrm{df}}, \qquad \frac{1}{\mu}\left|\sum_i \frac{\Delta x_i^k}{x_i^k}\right| \leq \epsilon_{\mathrm{gap}}$$

Long step Dikin's method

- long step given by $x^{k+1} = x^k + \tau \Delta x^k$
- $\tau = 0.95 \max\{t \mid x + t\Delta x \succeq 0\}$
- *i.e.*, step in Dikin direction, 95% of the way to boundary

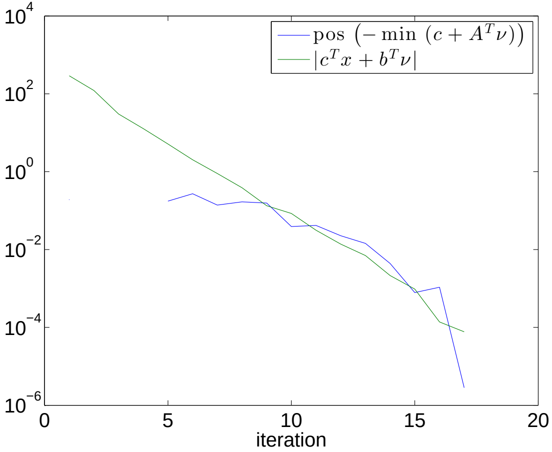
Short step example

$m = 100$, $n = 400$



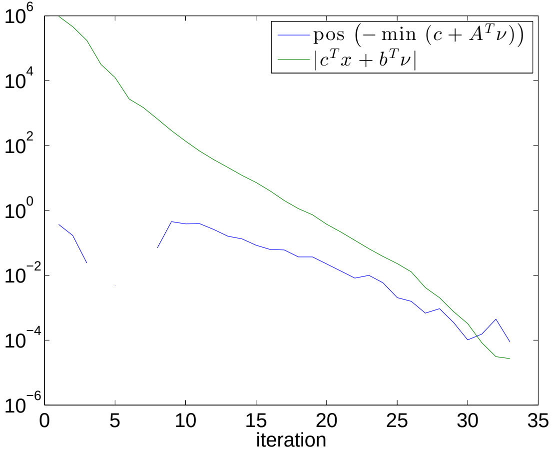
Long step example

$m = 100$, $n = 400$



Larger example

$m = 4000$, $n = 16000$, A 0.8% dense (500000 nonzeros), long step



Outline

Dikin's original method for LP

Generalized Dikin's method

Dikin ellipsoid for general constraints

- ϕ is self-concordant barrier for set $\mathcal{C}$
- for $y \in \textbf{int } \mathcal{C}$, $(x-y)^T \nabla^2 \phi(y)(x-y) \leq 1 \implies x \in \mathcal{C}$
- for $\mathcal{C} = \mathbf{R}^n_+$,

$$\phi(x) = -\sum_{i=1}^{n} \log(x_i), \quad \nabla^2 \phi(y) = \mathbf{diag}(y)^{-2}$$

- for $\mathcal{C} = \mathbf{S}^n_{++}$,

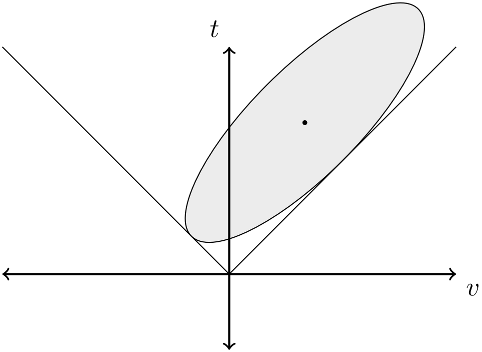
$$\phi(X) = -\log\det X, \quad \nabla^2 \phi(Y)(\Delta X) = Y^{-1}(\Delta X)Y^{-1}$$

- for direct product $\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_k$, use barrier

$$\phi(x) = \sum_{i=1}^{k} \phi_i(x_i)$$

SOCP Dikin ellipsoid

- $\mathcal{C} = \{x = (v, t) \in \mathbf{R}^n \times \mathbf{R} \mid \|v\|_2 \le t\}$
- $\phi(x) = -\log(t^2 - v^T v)$
- $\nabla^2 \phi(x) = \frac{2}{t^2 - v^T v} \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} + \frac{4}{(t^2 - v^T v)^2} \begin{bmatrix} -v \\ t \end{bmatrix} \begin{bmatrix} -v^T & t \end{bmatrix}$



Dikin's method with general constraints

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

- ϕ is self-concordant barrier for $\mathcal{C}$
- x^{k+1} is solution of

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & (x - x^k)^T H^k (x - x^k) \leq 1, \end{array}$$

  where $H^k = \nabla^2 \phi(x^k)$
- long step update: move 95% towards boundary

Inequality form LP

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

- $\phi(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$
- $H = \nabla^2 \phi(x) = A^T \mathbf{diag}(b - Ax)^{-2} A$
- Dikin step is

$$x^{k+1} = x^k - \frac{H^{-1}c}{\sqrt{c^T H^{-1} c}}$$

Inequality form LP cont.

- dual

$$\begin{array}{ll} \text{maximize} & -b^T\lambda \\ \text{subject to} & c + A^T\lambda = 0 \\ & \lambda \succeq 0 \end{array}$$

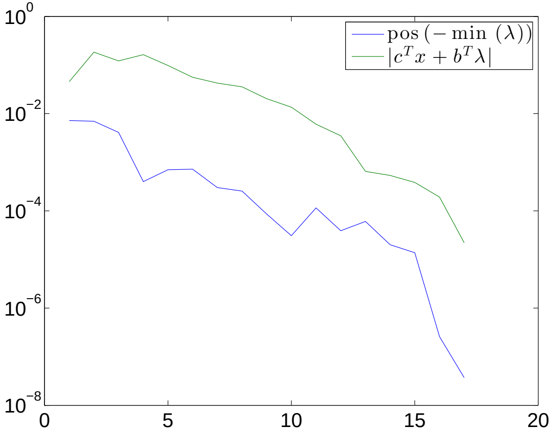
- stopping criteria

$$\lambda = \mathbf{diag}(b - Ax)^2 As \succeq -\epsilon_{\mathbf{df}}, \qquad |c^T x + b^T\lambda| \leq \epsilon_{\mathbf{gap}},$$

where $s = -H^{-1}c$

Inequality form LP example (long step)

$m = 1500$, $n = 500$



SDP

- SDP in inequality form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_{i=1}^n x_i A_i \preceq B \end{array}$$

  $B, A_i \in \mathbf{S}^m$

- $\phi(x) = -\log\det\left(B - \sum_{i=1}^n x_i A_i\right)$
- Hessian given by

$$H_{ij} = \mathbf{tr}(S^{-1} A_i S^{-1} A_j), \quad S = B - \sum_{i=1}^n x_i A_i$$

- step identical to inequality form LP

$$x^{k+1} = x^k - \frac{H^{-1} c}{\sqrt{c^T H^{-1} c}}$$

SDP cont.

- dual

$$\begin{array}{ll} \text{maximize} & -\mathbf{tr}(BZ) \\ \text{subject to} & c_i + \mathbf{tr}(A_i Z) = 0, \quad i = 1, \dots, n \\ & Z \succeq 0 \end{array}$$

- stopping criteria

$$Z \succeq -\epsilon_{\mathrm{df}} I, \qquad |c^T x + \mathbf{tr}(BZ)| \leq \epsilon_{\mathrm{gap}},$$

where $Z = \sum_{i=1}^{n} S^{-1} A_j S^{-1} s_j$, $s = -H^{-1} c$

SDP example (short step)

$m = 100$, $n = 100$