EE364a Review Session 8

- Algorithms review
- Example: Network rate optimization
- $\ell_{1.5}$ optimization

Optimality conditions

- Unconstrained convex problem with differentiable objective

$$\text{minimize} \quad f(x)$$

- Optimality condition: $\nabla f(x^\star) = 0$

- Convex problem with equality constraints, $A \in \mathbf{R}^{m \times n}$:

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

- Optimality conditions: $\exists \nu \in \mathbf{R}^m$

$$Ax^\star = b, \quad \nabla f(x^\star) + A^T \nu = 0$$

Inequality constrained problems

- Problems with inequality constraints:

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

- Exact reformulation with indicator function:

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

where $I_-(u) = 0$ if $u \leq 0$, $I_-(u) = \infty$ otherwise

- Approximation via *barrier* function:

$$\begin{array}{ll} \text{minimize} & tf_0(x) + \sum_{i=1}^m h(f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

- h convex (twice differentiable) increasing function with domain $-\mathbf{R}_{++}$

- Approximation improves as $t \to \infty$

- Logarithmic barrier function: $\phi(x) = -\sum_{i=1}^m \log(-f_i(x))$, with $\mathbf{dom}\,\phi = \{x \mid f_1(x) < 0, \ldots, f_m(x) < 0\}$

- Approximation via log-barrier:

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

is an equality constrained problem

- Example:

$$\begin{array}{ll} \text{minimize} & x^2+1 \\ \text{subject to} & 2 \le x \le 4, \end{array}$$

- Feasible set is $[2,4]$, and optimal point $x^\star = 2$

- Log-barrier function $\phi(x) = -\log(x-2) - \log(4-x)$

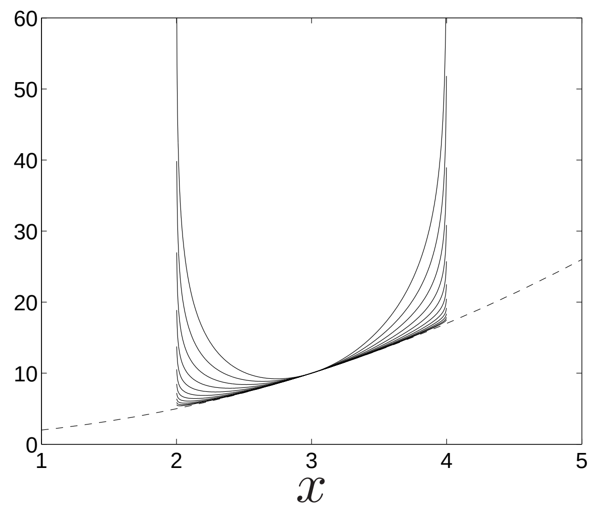

Figure 1 $f_0 + (1/t)\widehat{I}$, for $t = 10^{-1}$, $10^{-0.8}$, $10^{-0.6}$, . . . , $10^{0.8}$, 10.

Newton Method

- Newton step for the approximate problem:

$$\begin{bmatrix} t\nabla^2 f_0(x) + \nabla^2\phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\rm nt} \\ \nu_{\rm nt} \end{bmatrix} = - \begin{bmatrix} t\nabla f_0(x) + \nabla\phi(x) \\ 0 \end{bmatrix}$$

- Gradient and Hessian of the logarithmic barrier function ϕ are given by

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

Network rate optimization

- Graph with L edges, each with positive capacity c_i
- n flows with rates $x_1, \ldots, x_n \geq 0$ moving along predetermined paths in the network
- Total traffic on edge i: sum of flows passing through edge i
- Flow-edge incidence matrix $A \in \mathbf{R}^{L \times n}$

$$A_{ij} = \begin{cases} 1 & \text{flow } j \text{ passes through edge } i \\ 0 & \text{otherwise.} \end{cases}$$

- Capacity constraint on each edge: $(Ax)_i \leq c_i$

- Utility $U(x)$ associated with flow x

$$U(x) = U_1(x_1) + \cdots + U_n(x_n)$$

- U_i strictly concave, non-decreasing functions, with $\mathbf{dom}\, U_i = \mathbf{R}_{++}$

- Network rate optimization problem:

$$\begin{array}{ll} \text{maximize} & U(x) \\ \text{subject to} & Ax \preceq c, \quad x \succeq 0 \end{array}$$

Barrier method

- Barrier method: at each step, minimize

$$tU(x) - \sum_{i=1}^{L} \log(c - Ax)_i - \sum_{j=1}^{n} \log x_j,$$

  using Newton's method

- Newton step $\Delta x_{\rm nt}$ is solution to

$$(D_0 + A^T D_1 A + D_2)\Delta x_{\rm nt} = -g,$$

  where

$$\begin{aligned}
D_0 &= t\,\mathbf{diag}(U_1''(x), \ldots, U_n''(x)) \\
D_1 &= \mathbf{diag}(1/(c - Ax)_1^2, \ldots, 1/(c - Ax)_L^2) \\
D_2 &= \mathbf{diag}(1/x_1^2, \ldots, 1/x_n^2)
\end{aligned}$$

- Sparsity structure of coefficient matrix:

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

if and only if flow i and flow j share a edge

- if AA^T is sparse instead of $A^T A$, use matrix inversion lemma

$$\begin{aligned} \Delta x_{\mathrm{nt}} &= -(D_0 + D_2)^{-1} g + \\ &\quad (D_0 + D_2)^{-1} A^T (D_1^{-1} + A(D_0 + D_2)^{-1} A^T)^{-1} A (D_0 + D_2)^{-1} g \end{aligned}$$

- Sparsity pattern

$$(D_1^{-1} + A(D_0 + D_2)^{-1} A^T)_{ij} \neq 0$$

if and only if there is a flow that passes through edge i and edge j

Dual problem

- Conjugate utility functions $V_i = (-U_i)^*$:

$$V_i(\lambda) = \sup_{x>0}(\lambda x + U_i(x))$$

- Domain of V_i is $-R_{++}$
- Lagrangian

$$\begin{aligned} L(x,\lambda,\mu) &= \sum_{i=1}^{n}(-U_i(x)) + \lambda^T(Ax-c) - \mu^T x \\ &= -\sum_{i=1}^{n}\big(U_i(x) - (A^T\lambda)_i x_i + \mu_i x_i\big) - c^T\lambda. \end{aligned}$$

- Dual function $g(\lambda, \mu)$

$$\begin{aligned}\inf_x L(x, \lambda, \mu) &= -\sum_{i=1}^n \sup_x (U_i(x_i) - (A^T\lambda)_i x_i + \mu_i x_i) - c^T\lambda \\ &= -\sum_{i=1}^n V_i(-(A^T\lambda)_i + \mu_i) - c^T\lambda\end{aligned}$$

- Dual problem

$$\begin{array}{ll}\text{minimize} & c^T\lambda + \sum_{i=1}^n V_i(-(A^T\lambda)_i + \mu_i) \\ \text{subject to} & \lambda \succeq 0, \quad \mu \succeq 0\end{array}$$

- V_i is increasing on its domain ($-\mathbf{R}_{++}$), so $\mu = 0$ at the optimum:

$$\begin{array}{ll}\text{minimize} & c^T\lambda + \sum_{i=1}^n V_i(-(A^T\lambda)_i) \\ \text{subject to} & \lambda \succeq 0\end{array}$$

Barrier method for dual

- Hessian of $t\left(c^T\lambda + \sum_{i=1}^n V_i(-(A^T\lambda)_i)\right) - \sum_i \log \lambda_i$

$$H = tA\,\mathbf{diag}(V_i''(-(A^T\lambda)_i)A^T + \mathbf{diag}(\lambda)^{-2}$$

- If AA^T is sparse, solve the Newton equation $H\Delta\lambda = -g$
- If A^TA is sparse, we apply the matrix inversion lemma, etc, as in primal

$\ell_{1.5}$ optimization

$$\text{minimize} \quad \|Ax - b\|_{1.5} = \left(\sum_{i=1}^m |a_i^T x - b_i|^{3/2}\right)^{2/3}$$

problem:

1. give simple optimality conditions for this problem
2. formulate this problem as an SDP

Optimality conditions

equivalent problem

$$\text{minimize} \quad f(x) = \textstyle\sum_{i=1}^m |a_i^T x - b_i|^{3/2}$$

- objective differentiable
- use first order optimality conditions

$$\nabla f(x) = \sum_{i=1}^m (3/2)\,\mathbf{sign}(a_i^T x - b_i)|a_i^T x - b_i|^{1/2} a_i = 0$$

SDP formulation

equivalent problem

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T t \\ \text{subject to} & s^{3/2} \preceq t, \\ & -s_i \preceq a_i^T x - b_i \preceq s_i \quad i = 1, \ldots, m \end{array}$$

- variables $x \in \mathbf{R}^n$, $s, t \in \mathbf{R}^m$
- problem convex, but not an SDP
- need to transform $s^{3/2} \preceq t$ into an LMI

LMI transformation

- using

$$\begin{bmatrix} u & v \\ v & w \end{bmatrix} \succeq 0 \;\Leftrightarrow\; u \geq 0,\; uw \geq v^2$$

- we have that the constraint

$$s_i^{3/2} \leq t_i$$

- is equivalent to

$$\begin{bmatrix} \sqrt{s}_i & s_i \\ s_i & t \end{bmatrix} \succeq 0$$

- which in turn is equivalent to the LMI

$$\begin{bmatrix} y_i & s_i \\ s_i & t_i \end{bmatrix} \succeq 0, \quad \begin{bmatrix} s_i & y_i \\ y_i & 1 \end{bmatrix} \succeq 0$$

SDP formulation

putting it all together

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T t \\ \text{subject to} & -s_i \preceq a_i^T x - b_i \preceq s_i, \quad i = 1, \ldots, m \\ & \begin{bmatrix} y_i & s_i \\ s_i & t_i \end{bmatrix} \succeq 0, \quad \begin{bmatrix} s_i & y_i \\ y_i & 1 \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m \end{array}$$

- SDP with variables x, s, t, and y
- same technique can be used for other problems involving polynomials