Numerical Optimization

Exercise VI

Jonathan Balzer

Geometric Modeling and Scientific Visualization Center
King Abdullah University of Science and Technology

Course Overview

1. fundamentals
   - convergence rates
   - numerical differentiation (finite differencing, AD)
   - KKT condition
   - convexity
2. unconstrained optimization
   - gradient descent
   - Newton (Hessian modification) and Quasi-Newton methods
   - conjugate gradient method
3. line search
   - Wolfe/Armijo/Goldstein conditions
   - exact line search
   - interpolation

Course Overview

4. trust-region methods
   - Cauchy point
   - minimization under quadratic norm constraint
5. least-squares problems
   - linear
   - nonlinear
6. linear programming
   - simplex method
   - interior point method
7. quadratic programming
   - equality-constrained problems
   - active set method
   - interior point method
   - SQP

Course Overview

Briefly:

8. penalty methods

9. interior points methods for nonlinear programming

Exercises

1. There are K power plants $P_1 \ldots, P_K$ producing s_k gigawatts (GW) of electricity and L cities $C_1, \ldots, C_L$, each of which consumes r_l GW thereof. Let t_{kl} be the cost of transporting 1 GW of electricity from plant P_k to city C_l.

(i) Model the problem of how to meet the market requirements at minimal cost by a linear program.

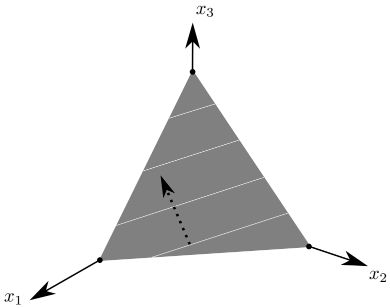
(ii) Bring it to standard form assuming

$$K = 2,\ s_1 = 2,\ s_2 = 1,\ L = 1,\ r_1 = 3,\ t_{11} = 3,\ t_{21} = 1.$$

Interior Point Methods

Idea:

- approach minimizer from *inside* the simplex
- solve nonlinear KKT system in $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s})$ by constrained Newton method for root-finding



Problem:

- boundary reached too quickly
- revert to active set method

Interior Point Methods

Why call it "primal-dual method"?

Recall:

$$F(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s}) := \begin{pmatrix} \mathbf{A}^\top \boldsymbol{\lambda} + \boldsymbol{s} - \boldsymbol{c} \\ \mathbf{A}\boldsymbol{x} - \boldsymbol{b} \\ \mathbf{XS}\boldsymbol{e} \end{pmatrix}$$

- $F(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s})$ involves primal and dual variables
- if all iterates *strictly* feasible, i.e.
  - $\mathbf{A}^\top \boldsymbol{\lambda}_k + \boldsymbol{s}_k - \boldsymbol{c} = \mathbf{0}$,
  - $\mathbf{A}\boldsymbol{x}_k - \boldsymbol{b} = \mathbf{0}$,
  - $\boldsymbol{x}_k > \mathbf{0}$, and $\boldsymbol{s}_k > \mathbf{0}$,

  then duality gap $\boldsymbol{x}^\top \boldsymbol{s} \to 0$ implies $F(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s}) \to \mathbf{0}$

Interior Point Methods

Newton equation for feasible point:

$$\begin{pmatrix} \mathbf{0} & \mathbf{A}^\top & \mathbf{I} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{S}_k & \mathbf{0} & \mathbf{X}_k \end{pmatrix} \begin{pmatrix} \triangle \boldsymbol{x}_k \\ \triangle \boldsymbol{\lambda}_k \\ \triangle \boldsymbol{s}_k \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ -\mathbf{X}_k \mathbf{S}_k \boldsymbol{e} \end{pmatrix}$$

$\Rightarrow \triangle \boldsymbol{x}_k \in \ker \mathbf{A}$ and $(\triangle \boldsymbol{\lambda}_k, \triangle \boldsymbol{s}_k) \in \ker \begin{pmatrix} \mathbf{A}^\top & \mathbf{I} \end{pmatrix}$

In other words: if $(\boldsymbol{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{s}_k)$
also $(\boldsymbol{x}_{k+1}, \boldsymbol{\lambda}_{k+1}, \boldsymbol{s}_{k+1})$ *primal-dual-feasible*,
but what about $\boldsymbol{x}, \boldsymbol{s} > \mathbf{0}$?

Interior Point Methods

- log-barrier modification:

$$\min_{\boldsymbol{x}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} - \tau \sum_{i=1}^{n} \ln x_i \right\}$$

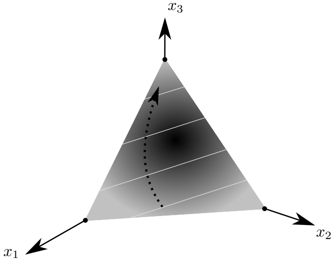
  subject to equality constraint

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$$

- zeros of

$$F_\tau(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s}) := \begin{pmatrix} \mathbf{A}^\top \boldsymbol{\lambda} + \boldsymbol{s} - \boldsymbol{c} \\ \mathbf{A}\boldsymbol{x} - \boldsymbol{b} \\ \mathbf{XS}\boldsymbol{e} - \tau \boldsymbol{e} \end{pmatrix}$$

  KKT points of augmented problem



Interior Point Methods

- obvious:

$$\lim_{\tau \to 0} F_\tau(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s}) \to F(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{s})$$

- choose $\tau = \sigma\mu$:
  - centering parameter $\sigma \in [0, 1]$
    - $\sigma = 0$: $F_\tau = F$
    - $\sigma = 1$: forces $x_1 s_1 = \ldots = x_n s_n$
  - duality gap

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i s_i \to 0$$

whenever $0 < \sigma < 1$