Lecture 8: Equality Constrained Minimization

Xiugang Wu

University of Delaware

Fall 2019

Outline

- Introduction
- Eliminating Equality Constraints
- Newton's Method With Equality Constraints
- Infeasible Start Newton Method

Outline

- Introduction
- Eliminating Equality Constraints
- Newton's Method With Equality Constraints
- Infeasible Start Newton Method

Equality Constrained Minimization

$$\text{minimize } f(x)$$
$$\text{subject to } Ax = b$$

- f convex, twice continuously differentiable
- $p^* = \inf_x f(x)$ is finite and attained
- $A \in \mathbf{R}^{p \times n}$ with rank p

Optimality condition: from KKT condition, x^* is optimal if and only if there exists ν^* such that

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

- a set of $n+p$ equations in $n+p$ variables x^*, ν^*

Equality Constrained Quadratic Minimization

$$\begin{aligned} \text{minimize } & (1/2)x^T P x + q^T x + r \\ \text{subject to } & Ax = b \end{aligned}$$

Optimality condition:

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

- coefficient matrix is called KKT matrix
- KKT matrix is nonsingular iff

$$Ax = 0, x \neq 0 \Rightarrow x^T P x > 0$$

- equivalent condition for nonsingulariry: $P + A^T A \succ 0$

Outline

- Introduction
- Eliminating Equality Constraints
- Newton's Method With Equality Constraints
- Infeasible Start Newton Method

Eliminating Equality Constraint

represent solution of $\{x|Ax = b\}$ as

$$\{x|Ax = b\} = \{Fz + \hat{x}|z \in \mathbf{R}^{n-p}\}$$

where $\hat{x}$ is any particular solution, and range of $F \in \mathbf{R}^{n\times(n-p)}$ is null space of A (F has rank $n - p$ and $AF = 0$)

reduced or eliminated problem

$$\text{minimize } f(Fz + \hat{x})$$

- an unconstrained problem with variable $z \in \mathbf{R}^{n-p}$
- once have solution z^*, can obtain x^* and ν^* as

$$x^* = Fz^* + \hat{x}, \;\; \nu^* = -(AA^T)^{-1}A\nabla f(x^*)$$

Example

optimal allocation with resource constraint

$$\begin{array}{ll} \text{minimize} & f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \\ \text{subject to} & x_1 + x_2 + \cdots + x_n = b \end{array}$$

eliminate $x_n = b - x_1 - \cdots - x_{n-1}$, i.e. choose $x = Fz + \hat{x}$ with

$$\hat{x} = be_n, \quad F = \begin{bmatrix} I \\ -\mathbf{1}^T \end{bmatrix}$$

reduced problem:

$$\text{minimize} \quad f_1(x_1) + f_2(x_2) + \cdots + f_n(b - x_1 - \cdots - x_{n-1})$$

Outline

- Introduction
- Eliminating Equality Constraints
- Newton's Method With Equality Constraints
- Infeasible Start Newton Method

Extension of Newton's Method

Newton's method can be extended to include equality constraints. The extended method is almost the same as Newton's method without constraints, except for two differences:

- The initial point must be feasible (i.e., satisfy $x \in \mathrm{dom}\ f$ and $Ax = b$)

- The definition of Newton step is modified to take the equality constraints into account. In particular, we make sure that the Newton step Δx_{nt} is a feasible direction, i.e., $A\Delta x_{\mathrm{nt}} = 0$

Newton Step

Newton step Δx_{nt} of f at feasible x is given by solution v of

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

Interpretations:
- Δx_{nt} minimizes second order approximation (with variable v)

$$\begin{array}{ll} \text{minimize} & \hat{f}(x+v) = f(x) + \nabla f(x)^T v + (1/2) v^T \nabla^2 f(x) v \\ \text{subject to} & A(x+v) = b \end{array}$$

- Δx_{nt} solves linearized optimality condition:

$$0 = \nabla f(x+v) + A^T w \approx \nabla f(x) + \nabla^2 f(x) v + A^T w, \;\; A(x+v) = b$$

Newton Decrement

$$\begin{aligned}\lambda(x) &= (\Delta x_{\mathrm{nt}}^T \nabla^2 f(x) \Delta x_{\mathrm{nt}})^{1/2} \\ &= \|\Delta x_{\mathrm{nt}}\|_{\nabla^2 f(x)}\end{aligned}$$

- gives an estimate of $f(x) - p^*$, using quadratic approximation $\hat{f}$:

$$f(x) - \inf_{Ay=b} \hat{f}(y) = \lambda(x)^2/2$$

- as before,

$$\nabla f(x)^T \Delta x_{\mathrm{nt}} = -\|\Delta x_{\mathrm{nt}}\|^2_{\nabla^2 f(x)} = -\lambda(x)^2$$

therefore it comes up in backtracking line search

- in general, $\lambda(x) \neq (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$

Newton's Method with Equality Constraints

---

given starting point $x \in \mathbf{dom}\, f$ with $Ax = b$, tolerance $\epsilon > 0$.

repeat

1. Compute the Newton step and decrement Δx_{nt}, $\lambda(x)$.
2. *Stopping criterion.* **quit** if $\lambda^2/2 \leq \epsilon$.
3. *Line search.* Choose step size t by backtracking line search.
4. *Update.* $x := x + t\Delta x_{\mathrm{nt}}$.

- a feasible descent method: $x^{(k)}$ feasible and $f(x^{k+1}) < f(x^k)$
- affine invariant

Newton's Method and Elimination

Newton's method for reduced problem

$$\text{minimize } \bar{f}(z) = f(Fz + \hat{x})$$

Newton's method for $\bar{f}$, started at $z^{(0)}$, generates iterates $z^{(k)}$

Newton's method with equality constraints when started at $x^{(0)} = Fz^{(0)} + \hat{x}$, iterates are

$$x^{(k+1)} = Fz^{(k)} + \hat{x}$$

hence, don't need separate convergence analysis

Outline

- Introduction
- Eliminating Equality Constraints
- Newton's Method With Equality Constraints
- Infeasible Start Newton Method

Newton Step at Infeasible Points

Linearizing optimality conditions at infeasible x gives

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\mathrm{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

Primal-dual Interpretations:
- write optimality condition as $r(y) = 0$, where

$$y = (x, \nu), \; r(y) = (\nabla f(x) + A^T \nu, Ax - b)$$

- linearizing $r(y) = 0$ gives $r(y + \Delta y) \approx r(y) + Dr(y)\Delta y = 0$:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\mathrm{nt}} \\ \Delta \nu_{\mathrm{nt}} \end{bmatrix} = - \begin{bmatrix} \nabla f(x) + A^T \nu \\ Ax - b \end{bmatrix}$$

Infeasible Start Newton Method

---

given starting point $x \in \mathbf{dom}\, f$, ν, tolerance $\epsilon > 0$, $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$.

repeat

1. Compute primal and dual Newton steps Δx_{nt}, $\Delta \nu_{\mathrm{nt}}$.
2. *Backtracking line search on* $\|r\|_2$.
   $t := 1$.
   while $\|r(x + t\Delta x_{\mathrm{nt}}, \nu + t\Delta \nu_{\mathrm{nt}})\|_2 > (1 - \alpha t)\|r(x, \nu)\|_2, \quad t := \beta t$.
3. *Update.* $x := x + t\Delta x_{\mathrm{nt}}$, $\nu := \nu + t\Delta \nu_{\mathrm{nt}}$.

until $Ax = b$ and $\|r(x, \nu)\|_2 \leq \epsilon$.

---

not a descent method: $f(x^+) > f(x)$ is possible

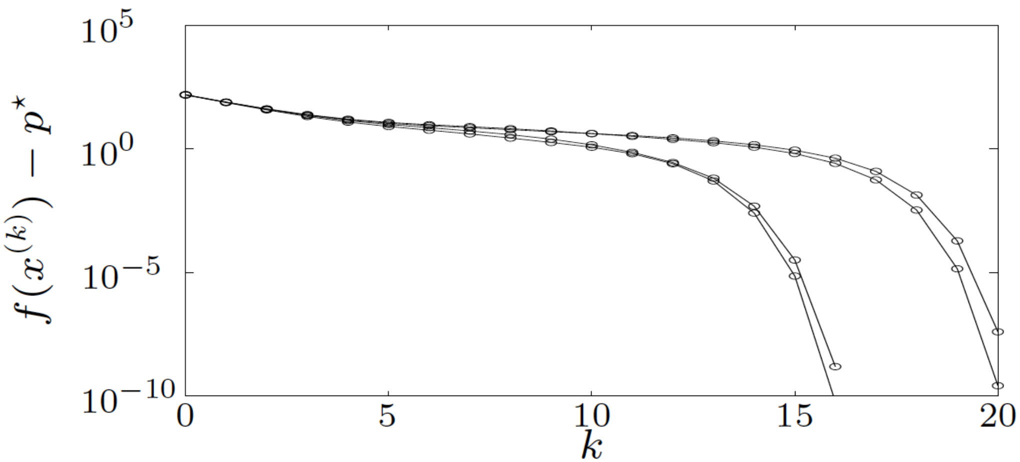
Equality Constrained Analytic Centering

$$\text{primal problem: minimize } -\sum_{i=1}^{n} \log x_i \text{ subject to } Ax = b$$

$$\text{dual problem: maximize } -b^T\nu + \sum_{i=1}^{n} \log(A^T\nu)_i + n$$

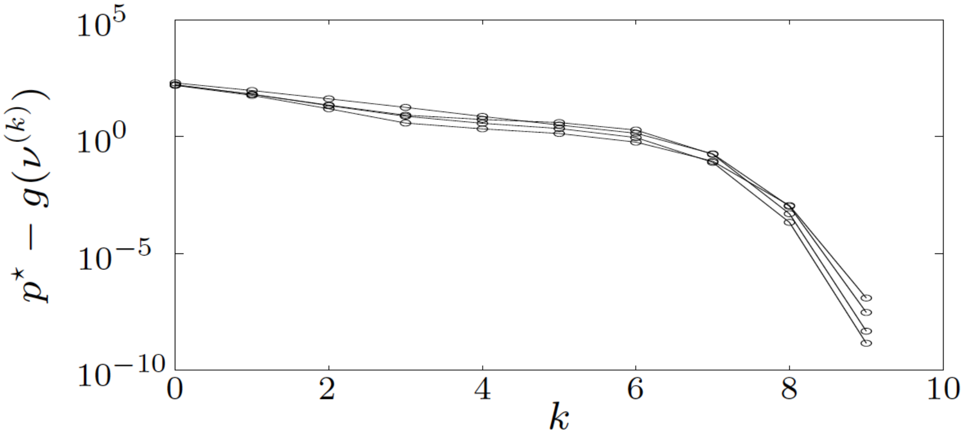
Three methods for an example with $A \in \mathbf{R}^{100\times 500}$, different starting points:

1. Newton method with equality constraints ($x^{(0)} \succ 0$, $Ax^{(0)} = b$)



Equality Constrained Analytic Centering

2. Newton method applied to dual problem $(A^T\nu^{(0)} \succ 0)$



3. infeasible start Newton method $(x^{(0)} \succ 0)$