Convex Optimization

Lecture 12 - Equality Constrained Optimization

Instructor: Yuanzhang Xiao

University of Hawaii at Manoa

Fall 2017

Today's Lecture

1. Basic Concepts

2. Newton's Methods for Equality Constrained Problems

Outline

1. Basic Concepts

2. Newton's Methods for Equality Constrained Problems

Equality Constrained Optimization Problems

equality constrained minimization problem:

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

- $f(x)$ convex, twice continuously differentiable
- $A \in \mathbb{R}^{p \times n}$ with $\textbf{rank} A = p < n$
- $p^\star = f(x^\star) = \inf \{f(x) \mid Ax = b\}$ attained and finite

optimality condition: there exists a $\nu^\star \in \mathbb{R}^p$ such that

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

Equality Constrained Quadratic Optimization

equality constrained quadratic minimization:

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

where $P \in \mathbb{S}^n_+$

optimality condition: there exists a $\nu^\star \in \mathbb{R}^p$ such that

$$Ax^\star = b, \quad Px^\star + q + A^T \nu^\star = 0$$

equivalent to

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

basis for Newton's method in equality constrained optimization

Solving Equality Constrained Optimization

how to solve equality constrained optimization?

- eliminate equality constraints
- solve dual problem and recover the solution to primal problem
- Newton's method for equality constrained optimization

Newton's method is most commonly used

- preserve structures (e.g., sparsity) of the problem

Eliminate Equality Constraints

find $F \in \mathbb{R}^{n\times(n-p)}$ and $\hat{x} \in \mathbb{R}^n$ such that

$$\{x \mid Ax = b\} = \left\{Fz + \hat{x} | z \in \mathbb{R}^{n-p}\right\}$$

- $\hat{x}$ is any particular solution to $Ax = b$
- F is any matrix such that $\textbf{range}(F) = \textbf{null}(A)$ (i.e., $AF = 0$)

the original problem equivalent to

$$\text{minimize} \quad \tilde{f}(z) = f\left(Fz + \hat{x}\right)$$

with optimization variable $z \in \mathbb{R}^{n-p}$

from $z^\star$, recover optimal primal and dual variables

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

choice F and $\hat{x}$ are not unique

Example – Optimal Allocation With Resource Constraints

resource allocation problem:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} f_i(x_i) \\ \text{subject to} & \sum_{i=1}^{n} x_i = b \end{array}$$

eliminating $x_n = b - \sum_{i=1}^{n-1} x_i$ is equivalent to

$$\hat{x} = be_n, \quad F = \begin{bmatrix} I \\ -\mathbf{1}^T \end{bmatrix} \in \mathbb{R}^{n \times (n-1)}$$

equivalent problem

$$\text{minimize} \quad \sum_{i=1}^{n-1} f_i(x_i) + f_n\left(b - \sum_{i=1}^{n-1} x_i\right)$$

Solve Dual Problem

Lagrangian:

$$L(x,\nu) = f(x) + \nu^T(Ax - b)$$

dual function:

$$\begin{aligned} g(\nu) &= \inf_x f(x) + \nu^T(Ax - b) \\ &= -b^T\nu + \inf_x \left(f(x) + (A^T\nu)^T x \right) \\ &= -b^T\nu - \sup \left(-f(x) - (A^T\nu)^T x \right) \\ &= -b^T\nu - f^*(-A^T\nu) \end{aligned}$$

where $f^*(y) \triangleq \sup_x \left(-f(x) + y^T x \right)$ is the conjugate of f

dual problem

$$\text{maximize} \quad -b^T\nu - f^*(-A^T\nu)$$

with optimization variable $\nu \in \mathbb{R}^p$

Example – Equality Constrained Analytic Center

equality constrained analytic center:

$$\begin{array}{ll} \text{minimize} & f(x) = -\sum_{i=1}^{n} \log x_i \\ \text{subject to} & Ax = b \end{array}$$

conjugate function ($\mathbf{dom} f = -\mathbb{R}^n_{++}$):

$$\begin{aligned} f^*(y) &\triangleq \sup_x \left(-f(x) + y^T x\right) \\ &= \sup_x \sum_{i=1}^{n} \log x_i + y^T x \\ &= \sup_x \sum_{i=1}^{n} (\log x_i + y_i x_i) \\ &= -n - \textstyle\sum_{i=1}^{n} \log(-y_i) \end{aligned}$$

Example – Equality Constrained Analytic Center

dual problem:

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

with implicit constraints $A^T\nu > 0$

how to reconstruct primal solution from dual solution?

optimality condition:

$$\nabla f(x^\star) + A^T\nu^\star = -(1/x_1^\star, \ldots, 1/x_n^\star)^T + A^T\nu^\star = 0$$

which implies

$$x_i^\star = 1/(A^T\nu^\star)_i, \;\; i = 1, \ldots, n$$

Outline

1. Basic Concepts

2. Newton's Methods for Equality Constrained Problems

Newton's Methods

assume that the initial point is feasible

$$Ax^{(0)} = b$$

choose Newton step Δx_{nt} such that $x + \Delta x_{\text{nt}}$ is feasible

second-order Taylor approximation:

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

optimality condition for equality constrained quadratic problem:

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

Newton's Method For Equality Constrained Optimization

Newton's method for equality constrained optimization:

- a starting point $x \in \mathbf{dom} f$ with $Ax = b$
- repeat the following steps
  1. Newton step Δx_{nt} and Newton decrement

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

  2. quit if $\frac{\lambda^2}{2} \leq \epsilon$
  3. exact or backtracking line search
  4. update $x := x + t\Delta x_{\text{nt}}$

require a feasible starting point ("feasible descent method")

convergence results basically the same as unconstrained cases

Infeasible Start Newton's Methods

second-order Taylor approximation (x may be infeasible):

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

solving for Newton step:

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

optimality condition:

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

Interpretation as Primal-Dual Newton Step

recall optimality conditions:

$$\underbrace{Ax^\star - b = 0}_{\text{primal feasibility}}, \quad \underbrace{\nabla f(x^\star) + A^T \nu^\star = 0}_{\text{dual feasibility}}$$

define residual

$$r(x,\nu) = (r_{\text{dual}}(x,\nu), r_{\text{pri}}(x, nu)) = (\nabla f(x) + A^T\nu, Ax - b) \in \mathbb{R}^n \times \mathbb{R}^p$$

first-order Taylor approximation of residual $r(x,\nu)$:

$$r(x+\Delta x, \nu + \Delta\nu) \approx r(x,\nu) + Dr(x,\nu) \begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix}$$

where $Dr(x,\nu) \in \mathbb{R}^{(n+p)\times(n+p)}$ is the derivative of $r(x,\nu)$

trying to make the residual equal to zero:

$$Dr(x,\nu) \begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix} = -r(x,\nu)$$

Interpretation as Primal-Dual Newton Step

trying to make the residual equal to zero:

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

writing $\nu^+ = \nu + \Delta \nu$, we have

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

same as the optimality condition with

$$\Delta x_{\text{nt}} = \Delta x, \quad w = \nu^+ = \nu + \Delta \nu$$

Features of Infeasible Start Newton Method

objective value may not decrease:

it is possible that $f(x + t\Delta x) \geq f(x), \ \forall t \geq 0$

residual always decreases:

$$\|r(x + t\Delta x, \nu + t\Delta \nu)\|_2 < \|r(x, \nu)\|_2 \text{ for some } t > 0$$

$\Rightarrow$ line search based on $\|r\|_2$

full step feasibility:

$x + \Delta x$ is feasible

Infeasible Start Newton Method

infeasible start Newton method for equality constrained optimization:

- a starting point $x \in \mathbf{dom} f$, ν
- repeat the following steps
  1. compute primal and dual Newton steps Δx_{nt} and $\Delta \nu_{\text{nt}}$
  2. backtracking line search on $\|r\|_2$
     1. $t := 1$
     2. while $\|r(x + t\Delta x_{\text{nt}}, \nu + t\Delta \nu_{\text{nt}})\|_2 > (1 - \alpha t)\|r(x, \nu)\|_2$, $t := \beta t$
  3. update $x := x + t\Delta x_{\text{nt}}$ and $\nu := \nu + t\Delta \nu_{\text{nt}}$
- until $Ax = b$ and $\|r(x, \nu)\|_2 < \epsilon$