15.093 Optimization Methods

Lecture 17: Applications of Nonlinear Optimization

1 Lecture Outline

Slide 1

- History of Nonlinear Optimization
- Where do NLPs Arise?
- Portfolio Optimization
- Traffic Assignment
- The general problem
- The role of convexity
- Convex optimization
- Examples of convex optimization problems

2 History of Optimization

Slide 2

Fermat, 1638; Newton, 1670

$$\min \; f(x) \qquad \text{x: scalar}$$

$$\frac{df(x)}{dx} = 0$$

Euler, 1755

$$\min \; f(x_1, \ldots, x_n)$$

$$\nabla f(\boldsymbol{x}) = 0$$

Slide 3

Lagrange, 1797

$$\begin{aligned} \min \quad & f(x_1, \ldots, x_n) \\ \text{s.t.} \quad & g_k(x_1, \ldots, x_n) = 0 \qquad k = 1, \ldots, m \end{aligned}$$

Euler, Lagrange Problems in infinite dimensions, calculus of variations.

Kuhn and Tucker, 1950s Optimality conditions.

1950s Applications.

1960s Large Scale Optimization.

Karmakar, 1984 Interior point algorithms.

3 Where do NLPs Arise?

3.1 Wide Applicability

SLIDE 4

- Transportation

  Traffic management, Traffic equilibrium ...

  Revenue management and Pricing

- Finance - Portfolio Management
- Equilibrium Problems

SLIDE 5

- Engineering

  Data Networks and Routing

  Pattern Classification

- Manufacturing

  Resource Allocation

  Production Planning

4 A Simple Portfolio Selection Problem

4.1 Data

SLIDE 6

- x_i: decision variable on amount to invest in stock $i = 1, 2$
- r_i: reward from stock $i = 1, 2$ (random variable)
  Data:
- $\mu_i = E(r_i)$: expected reward from stock $i = 1, 2$
- $Var(r_i)$: variance in reward from stock $i = 1, 2$
- $\sigma_{ij} = E[(r_j - \mu_j)(r_i - \mu_i)] = Cov(r_i, r_j)$
- Budget B, target β on expected portfolio reward

5 A Simple Portfolio Selection Problem

5.1 The Problem

SLIDE 7

Objective: Minimize total portfolio variance so that:

- Expected reward of total portfolio is above target β
- Total amount invested stay within our budget
- No short sales

SLIDE 8

$$\min \ f(x) = x_1^2 Var(r_1) + x_2^2 Var(r_2) + 2x_1 x_2 \sigma_{12}$$

subject to

$$\sum_i x_i \le B$$

$$E[\sum_i r_i x_i] = \sum_i \mu_i x_i \ge \beta, \quad \text{(exp reward of portf.)}$$

$$x_i \ge 0, \quad i = 1, 2$$

(Linearly constrained NLP)

6 A Real Portfolio Optimization Problem

6.1 Data

SLIDE 9

- We currently own z_i shares from stock i, $i \in S$
- P_i: current price of stock i
- We consider buying and selling stocks in S, and consider buying new stocks from a set B ($B \cap S = \emptyset$)
- Set of stocks $B \cup S = \{1, \ldots, n\}$

SLIDE 10

- Data: Forecasted prices next period (say next month) and their correlations:

$$E[\hat{P}_i] = \mu_i$$

$$Cov(\hat{P}_i, \hat{P}_j) = E[(\hat{P}_i - \mu_i)(\hat{P}_j - \mu_j)] = \sigma_{ij}$$

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)', \qquad \boldsymbol{\Sigma} = \sigma_{ij}$$

6.2 Issues and Objectives

SLIDE 11

- Mutual funds regulations: we cannot sell a stock if we do not own it
- Transaction costs
- Turnover
- Liquidity
- Volatility
- **Objective:** Maximize expected wealth next period minus transaction costs

6.3 Decision variables

SLIDE 12

$$x_i = \begin{cases} \#\text{ shares bought or sold} & \text{if } i \in S \\ \#\text{ shares bought} & \text{if } i \in B \end{cases}$$

By convention:

$$\begin{array}{ll} x_i \geq 0 & \text{buy} \\ x_i < 0 & \text{sell} \end{array}$$

6.4 Transaction costs

SLIDE 13

- Small investors only pay commision cost: a_i \$/share traded
- Transaction cost: $a_i|x_i|$
- Large investors (like portfolio managers of large funds) may affect price: price becomes $P_i + b_i x_i$
- Price impact cost: $(P_i + b_i x_i)x_i - P_i x_i = b_i x_i^2$
- Total cost model:

$$c_i(x_i) = a_i|x_i| + b_i x_i^2$$

6.5 Liquidity

SLIDE 14

- Suppose you own 50% of all outstanding stock of a company
- How difficult is to sell it?
- Reasonable to bound the percentage of ownership on a particular stock
- Thus, for **liquidity** reasons $\dfrac{z_i + x_i}{z_i^{total}} \leq \gamma_i$
- z_i^{total} =# outstanding shares of stock i
- γ_i maximum allowable percentage of ownership

6.6 Turnover

SLIDE 15

- Because of transaction costs: $|x_i|$ should be small

$$|x_i| \leq \delta_i \quad \Rightarrow \quad -\delta_i \leq x_i \leq \delta_i$$

- Alternatively, we might want to bound turnover:

$$\sum_{i=1}^{n} P_i |x_i| \leq t$$

6.7 Balanced portfolios

SLIDE 16

- Need the value of stocks we buy and sell to balance out:

$$\left| \sum_{i=1}^{n} P_i x_i \right| \leq L \quad \Rightarrow \quad -L \leq \sum_{i=1}^{n} P_i x_i \leq L$$

- No short sales:

$$z_i + x_i \geq 0, \qquad i \in B \cup S$$

6.8 Expected value and Volatility

SLIDE 17

- Expected value of portfolio:

$$E\left[\sum_{i=1}^{n} \hat{P}_i (z_i + x_i)\right] = \sum_{i=1}^{n} \mu_i (z_i + x_i)$$

- Variance of the value of the portfolio:

$$Var\left[\sum_{i=1}^{n} \hat{P}_i (z_i + x_i)\right] = (\boldsymbol{z} + \boldsymbol{x})' \boldsymbol{\Sigma} (\boldsymbol{z} + \boldsymbol{x})$$

6.9 Overall formulation

SLIDE 18

$$\begin{aligned}
\max \quad & \sum_{i=1}^{n} \mu_i (z_i + x_i) - \sum_{i=1}^{n} (a_i |x_i| + b_i x_i^2) \\
\text{s.t.} \quad & (\boldsymbol{z} + \boldsymbol{x})' \boldsymbol{\Sigma} (\boldsymbol{z} + \boldsymbol{x}) \leq \boldsymbol{\sigma}^2 \\
& z_i + x_i \leq \gamma_i z_i^{total} \\
& -\delta_i \leq x_i \leq \delta_i \\
& -L \leq \sum_{i=1}^{n} P_i x_i \leq L \\
& \sum_{i=1}^{n} P_i |x_i| \leq t \\
& z_i + x_i \geq 0
\end{aligned}$$

7 The general problem

SLIDE 19

$$f(\boldsymbol{x}):\ \Re^n \mapsto \Re$$
$$g_i(\boldsymbol{x}):\ \Re^n \mapsto \Re, i = 1, \ldots, m$$

$$\boxed{\begin{array}{llcll} NLP: & \min & f(\boldsymbol{x}) & & \\ & \text{s.t.} & g_1(\boldsymbol{x}) & \leq & 0 \\ & & \vdots & & \\ & & g_m(\boldsymbol{x}) & \leq & 0 \end{array}}$$

7.1 Is Portfolio Optimization an NLP?

SLIDE 20

$$\begin{array}{ll} \max & \displaystyle\sum_{i=1}^{n} \mu_i(z_i + x_i) - \sum_{i=1}^{n}(a_i|x_i| + b_i x_i^2) \\ \text{s.t.} & (\boldsymbol{z}+\boldsymbol{x})'\boldsymbol{\Sigma}(\boldsymbol{z}+\boldsymbol{x}) \leq \sigma^2 \\ & z_i + x_i \leq \gamma_i z_i^{total} \\ & -\delta_i \leq x_i \leq \delta_i \\ & -L \leq \displaystyle\sum_{i=1}^{n} P_i x_i \leq L \\ & \displaystyle\sum_{i=1}^{n} P_i|x_i| \leq t \\ & z_i + x_i \geq 0 \end{array}$$

8 Geometry Problems

8.1 Fermat-Weber Problem

SLIDE 21

Given m points $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \in \Re^n$ (locations of retail outlets) and weights $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m \in \Re$. Choose the location of a distribution center.
That is, the point $\boldsymbol{x} \in \Re^n$ to minimize the sum of the weighted distances from $\boldsymbol{x}$ to each of the points $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \in \Re^n$ (**minimize total daily distance traveled**).

$$\begin{array}{ll} \min & \sum_{i=1}^{m} \boldsymbol{w}_i||\boldsymbol{x} - \boldsymbol{c}_i|| \\ \text{s.t.} & \boldsymbol{x} \in \Re^n \end{array}$$

or

$$
\begin{array}{ll}
\min & \sum_{i=1}^{m} \boldsymbol{w}_i ||\boldsymbol{x} - \boldsymbol{c}_i|| \\
\text{s.t.} & \boldsymbol{x} \geq 0 \\
 & A\boldsymbol{x} \leq b, \quad \text{feasible sites}
\end{array}
$$

(Linearly constrained NLP)

8.2 The Ball Circumscription Problem

SLIDE 22

Given m points $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \in \Re^n$, locate a distribution center at point $\boldsymbol{x} \in \Re^n$ to minimize the maximum distance from $\boldsymbol{x}$ to any of the points $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \in \Re^n$.

$$
\begin{array}{ll}
\min & \delta \\
\text{s.t.} & ||\boldsymbol{x} - \boldsymbol{c}_i|| \leq \delta, \qquad i = 1, \ldots, m
\end{array}
$$

9 Transportation

9.1 Traffic Assignment

SLIDE 23

- OD w, paths $p \in P_w$, demand d_w, x_p: flow of p
  $c_{ij}(\sum_{p:\ crossing\ (i,j)} x_p)$: travel cost of link (i, j).
  $c_p(x)$ is the travel cost of path p and

  $$c_p(x) = \sum_{(i,j)\ on\ p} c_{ij}(x_{ij}), \quad \forall p \in P_w, \quad \forall w \in W.$$

  System – optimization principle : Assign flow on each path to satisfy total demand and so that the total network cost is minimized.

$$Min\ C(x) = \sum_p c_p(x) x_p$$

$$\text{s.t.} \ \ x_p \geq 0, \quad \sum_{p \in P_w} x_p = d_w, \quad \forall w$$

9.2 Example

SLIDE 24

Consider a three path network, $d_w = 10$.
With travel costs $c_{p_1}(x) = 2x_{p_1} + x_{p_2} + 15$, $c_{p_2}(x) = 3x_{p_2} + x_{p_1} + 11$ $c_{p_3}(x) = x_{p_3} + 48$

$$C(x) = c_{p1}(x)x_{p1} + c_{p2}(x)x_{p2} + c_{p3}(x)x_{p3} =$$

$$2x_{p1}^2 + 3x_{p2}^2 + x_{p3}^2 + 2x_{p1}x_{p2} + 15x_{p1} + 11x_{p2} + 38x_{p3}$$

$$x_{p1}^* = 6, \quad x_{p2}^* = 4, \quad x_{p3}^* = 0$$

SLIDE 25

- **User – optimization principle** : Each user of the network chooses, among all paths, a path requiring minimum travel cost,
  i.e., for all $w \in W$ and $p \in P_w$,

$$x_p^* > 0 : \quad \longrightarrow \quad c_p(x^*) \leq c_{p'}(x^*) \quad \forall p' \in P_w, \quad \forall w \in W$$

  where $c_p(x)$ is the travel time of path p and

$$c_p(x) = \sum_{(i,j)\ on\ p} c_{ij}(x_{ij}), \quad \forall p \in P_w, \quad \forall w \in W$$

10 Optimal Routing

Slide 26

• Given a data net and a set W of OD pairs $w = (i, j)$ each OD pair w has input traffic d_w
• Optimal routing problem:

$$Min \quad C(x) = \sum_{i,j} C_{i,j}(\sum_{p:\ (i,j) \in p} x_p)$$

$$s.t. \quad \sum_{p \in P_w} x_p = d_w, \quad \forall w \in W$$

$$x_p \geq 0, \quad \forall p \in P_w, \quad w \in W$$

11 The general problem again

Slide 27

$$f(\boldsymbol{x}): \ \Re^n \mapsto \Re$$

is a continuous (usually differentiable) function of n variables

$$g_i(\boldsymbol{x}): \ \Re^n \mapsto \Re, i = 1, \ldots, m,$$

$$h_j(\boldsymbol{x}): \ \Re^n \mapsto \Re, j = 1, \ldots, l$$

$$\boxed{\begin{array}{llrcl} NLP: & \min & f(\boldsymbol{x}) & & \\ & \text{s.t.} & g_1(\boldsymbol{x}) & \leq & 0 \\ & & \vdots & & \\ & & g_m(\boldsymbol{x}) & \leq & 0 \\ & & h_1(\boldsymbol{x}) & = & 0 \\ & & \vdots & & \\ & & h_l(\boldsymbol{x}) & = & 0 \end{array}}$$

11.1 Definitions

Slide 28

- The feasible region of $NLOP$ is the set:

$$\mathcal{F} = \{x | g_1(\boldsymbol{x}) \leq 0, \ldots, g_m(\boldsymbol{x}) \leq 0\}$$
$$h_1(\boldsymbol{x}) = 0, \ldots, h_l(\boldsymbol{x}) = 0\}$$

11.2 Where do optimal solutions lie?

Slide 29

Example:

$$\min \ f(x, y) = (x - a)^2 + (y - b)^2$$

Subject to

$$(x - 8)^2 + (y - 9)^2 \leq 49$$
$$2 \leq x \leq 13$$
$$x + y \leq 24$$

Optimal solution(s) **do not necessarily** lie at an extreme point!
Depends on (a, b).

$(a, b) = (16, 14)$ then solution lies at a corner
$(a, b) = (11, 10)$ then solution lies in interior
$(a, b) = (14, 14)$ then solution lies on the boundary
(not necessarily corner)

11.3 Local vs Global Minima

Slide 30

- The ball centered at $\bar{\boldsymbol{x}}$ with radius ϵ is the set:

$$B(\bar{\boldsymbol{x}}, \epsilon) := \{\boldsymbol{x} | ||\boldsymbol{x} - \bar{\boldsymbol{x}}|| \leq \epsilon\}$$

- $\boldsymbol{x} \in \mathcal{F}$ is a *local minimum* of $NLOP$ if there exists $\epsilon > 0$ such that $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in B(\boldsymbol{x}, \epsilon) \cap \mathcal{F}$

- $\boldsymbol{x} \in \mathcal{F}$ is a *global minimum* of $NLOP$ if $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{F}$

12 Convex Sets

Slide 31

- A subset $S \subset \Re^n$ is a *convex set* if

$$\boldsymbol{x}, \boldsymbol{y} \in S \Rightarrow \lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in S \qquad \forall \lambda \in [0, 1]$$

- If S, T are convex sets, then $S \cap T$ is a convex set
- Implication: The intersection of any collection of convex sets is a convex set

13 Convex Functions

SLIDE 32

- A function $f(\boldsymbol{x})$ is a *convex function* if

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

$$\forall \boldsymbol{x}, \boldsymbol{y} \qquad \forall \lambda \in [0,1]$$

- A function $f(\boldsymbol{x})$ is a *concave function* if

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

$$\forall \boldsymbol{x}, \boldsymbol{y} \qquad \forall \lambda \in [0,1]$$

13.1 Examples in one dimension

SLIDE 33

- $f(x) = ax + b$
- $f(x) = x^2 + bx + c$
- $f(x) = |x|$
- $f(x) = -\ln(x)$ for $x > 0$
- $f(x) = \frac{1}{x}$ for $x > 0$
- $f(x) = e^x$

13.2 Properties

SLIDE 34

- If $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ are convex functions, and $a, b \geq 0$, then $f(\boldsymbol{x}) := af_1(\boldsymbol{x}) + bf_2(\boldsymbol{x})$ is a convex function
- If $f(\boldsymbol{x})$ is a convex function and $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}$, then $g(\boldsymbol{y}) := f(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})$ is a convex function

13.3 Recognition of a Convex Function

SLIDE 35

A function $f(\boldsymbol{x})$ is *twice differentiable* at $\bar{\boldsymbol{x}}$ if there exists a vector $\nabla f(\bar{\boldsymbol{x}})$ (called the *gradient of* $f(\cdot)$) and a symmetric matrix $H(\bar{\boldsymbol{x}})$ (called the *Hessian of* $f(\cdot)$) for which:

$$\begin{aligned} f(\boldsymbol{x}) &= f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})'(\boldsymbol{x} - \bar{\boldsymbol{x}}) \\ &\quad + \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})' H(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) + R(\boldsymbol{x})||\boldsymbol{x} - \bar{\boldsymbol{x}}||^2 \end{aligned}$$

where $R(\boldsymbol{x}) \to 0$ as $\boldsymbol{x} \to \bar{\boldsymbol{x}}$ SLIDE 36

The gradient vector is the vector of partial derivatives:

$$\nabla f(\bar{\boldsymbol{x}}) = \left(\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_1}, \ldots, \frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_n} \right)'$$

The Hessian matrix is the matrix of second partial derivatives:

$$H(\bar{\boldsymbol{x}})_{ij} = \frac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_i \partial x_j}$$

13.4 Examples

SLIDE 37

- For LP, $f(x) = c'x$, $\nabla f(\bar{x}) = c$
- For NLP,

$f(x) = 8x_1^2 - x_1x_2 + x_2^2 + 8x_1$, at $\bar{x} = (1, 0)$,
$f(\bar{x}) = 16$ and
$\nabla f(\bar{x})' = (16\bar{x}_1 - \bar{x}_2 + 8, -\bar{x}_1 + 2\bar{x}_2) = (24, -1)$.

$$\boldsymbol{H}(\bar{\boldsymbol{x}}) = \begin{bmatrix} 16 & -1 \\ -1 & 2 \end{bmatrix}$$

SLIDE 38

Property: $f(\boldsymbol{x})$ is a convex function if and only if $H(\boldsymbol{x})$ is positive semi-definite (PSD) for all $\boldsymbol{x}$

Recall that $\boldsymbol{A}$ is PSD if $\boldsymbol{u'Au} \geq 0, \quad \forall u$

Property: If $H(\boldsymbol{x})$ is PD for all $\boldsymbol{x}$, then $f(\boldsymbol{x})$ is a strictly convex function

13.5 Examples in n Dimensions

SLIDE 39

- $f(\boldsymbol{x}) = \boldsymbol{a'x} + b$
- $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x'Mx} - \boldsymbol{c'x}$ where $\boldsymbol{M}$ is *PSD*
- $f(\boldsymbol{x}) = ||\boldsymbol{x}||$ for any norm $||\cdot||$
- $f(\boldsymbol{x}) = \sum_{i=1}^{m} -\ln(b_i - \boldsymbol{a_i'x})$ for $\boldsymbol{x}$ satisfying $\boldsymbol{Ax} < \boldsymbol{b}$

14 Convex Optimization

14.1 Convexity and Minima

SLIDE 40

$$\begin{array}{ll} \min & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{F} \end{array}$$

Theorem: Suppose that $\mathcal{F}$ is a convex set, $f : \mathcal{F} \to \Re$ is a convex function, and $\boldsymbol{x}^*$ is a local minimum of P. Then $\boldsymbol{x}^*$ is a global minimum of f over $\mathcal{F}$.

14.1.1 Proof

SLIDE 41

Assume that $\boldsymbol{x}^*$ is not the global minimum. Let $\boldsymbol{y}$ be the global minimum. From the convexity of $f(\cdot)$,

$$f(\boldsymbol{y}(\lambda)) = f(\lambda\boldsymbol{x}^* + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}^*) + (1-\lambda)f(\boldsymbol{y}) < \lambda f(\boldsymbol{x}^*) + (1-\lambda)f(\boldsymbol{x}^*) = f(\boldsymbol{x}^*)$$

for all $\lambda \in (0,1)$.

Therefore, $f(\boldsymbol{y}(\lambda)) < f(\boldsymbol{x}^*)$ for all $\lambda \in (0,1)$, and so $\boldsymbol{x}^*$ is not a local minimum, resulting in a contradiction

14.2 COP

SLIDE 42

$$
\begin{array}{rll}
COP: & \min & f(\boldsymbol{x}) \\
& \text{s.t.} & g_1(\boldsymbol{x}) \le 0 \\
& & \vdots \\
& & g_m(\boldsymbol{x}) \le 0 \\
& & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{array}
$$

SLIDE 43

COP is called a *convex optimization problem* if $f(\boldsymbol{x}), g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x})$ are convex functions

Note that this implies that the feasible region $\mathcal{F}$ is a convex set

In COP we are minimizing a convex function over a convex set

Implication: If COP is a convex optimization problem, then any local minimum will be a global minimum.

15 Examples of COPs

SLIDE 44

The Fermat-Weber Problem - COP

$$
\begin{array}{ll}
\min & \sum\limits_{i=1}^{m} \boldsymbol{w}_i ||\boldsymbol{x} - \boldsymbol{c}_i|| \\
\text{s.t.} & \boldsymbol{x} \in P
\end{array}
$$

The Ball Circumscription Problem - COP

$$
\begin{array}{lll}
\min & \delta & \\
\text{s.t.} & ||\boldsymbol{x} - \boldsymbol{c}_i|| \le \delta, & i = 1, \ldots, m
\end{array}
$$

15.1 Is Portfolio Optimization a COP?

SLIDE 45

$$\begin{array}{ll} \max & \sum_{i=1}^{n} \mu_i(z_i + x_i) - \sum_{i=1}^{n} (a_i|x_i| + b_i x_i^2) \\ \text{s.t.} & (\boldsymbol{z} + \boldsymbol{x})' \boldsymbol{\Sigma} (\boldsymbol{z} + \boldsymbol{x}) \leq \sigma^2 \\ & z_i + x_i \leq \gamma_i z_i^{total} \\ & -\delta_i \leq x_i \leq \delta_i \\ & -L \leq \sum_{i=1}^{n} P_i x_i \leq L \\ & \sum_{i=1}^{n} P_i |x_i| \leq t \\ & z_i + x_i \geq 0 \end{array}$$

15.2 Quadratically Constrained Problems

SLIDE 46

$$\begin{array}{lrcl} \min & (\boldsymbol{A_0 x} + \boldsymbol{b_0})'(\boldsymbol{A_0 x} + \boldsymbol{b_0}) - \boldsymbol{c_0' x} - d_0 & & \\ \text{s.t.} & (\boldsymbol{A_i x} + \boldsymbol{b_i})'(\boldsymbol{A_i x} + \boldsymbol{b_i}) - \boldsymbol{c_i' x} - d_i & \leq & 0 \\ & i = 1, \ldots, m & & \end{array}$$

This is a COP

16 Classification of NLPs

SLIDE 47

- **Linear**: $f(x) = c^t x$, $g_i(x) = A_i^t x - b_i$, $i = 1, ..., m$
- **Unconstrained**: $f(x)$, $\Re^n$
- **Quadratic**: $f(x) = c^t x + x^t Q x$, $g_i(x) = A_i^t x - b_i$
- **Linearly Constrained**: $g_i(x) = A_i^t x - b_i$
- **Quadratically Constrained**: $g_i(x) = (\boldsymbol{A_i x} + \boldsymbol{b_i})'(\boldsymbol{A_i x} + \boldsymbol{b_i}) - \boldsymbol{c_i' x} - d_i \leq 0$,
  $i = 1, \ldots, m$
- **Separable:** $f(x) = \sum_j f_j(x_j)$, $g_i(x) = \sum_j g_{ij}(x_j)$

17 Two Main Issues

SLIDE 48

- Characterization of minima

  Necessary — Sufficient Conditions

  Lagrange Multiplier and KKT Theory

- Computation of minima via iterative algorithms

  Iterative descent Methods

  Interior Point Methods

18 Summary

- Convex optimization is a powerful modeling framework
- Main message: convex optimization can be solved efficiently

15.093J / 6.255J Optimization Methods
Fall 2009