Systems Engineering Institute
Ministry of Education Key Lab for Intelligent Networks and Network Security

约束最优化问题
Constrained optimization problem

电信学院·自动化科学与技术系
系统工程研究所
吴江

Outline

- 方法概述
- 最优性条件

约束优化问题

$$
\begin{aligned}
\min \quad & z = f(x) \\
s.t. \quad & g_i(x) \le 0 \\
& h_j(x) = 0
\end{aligned}
$$

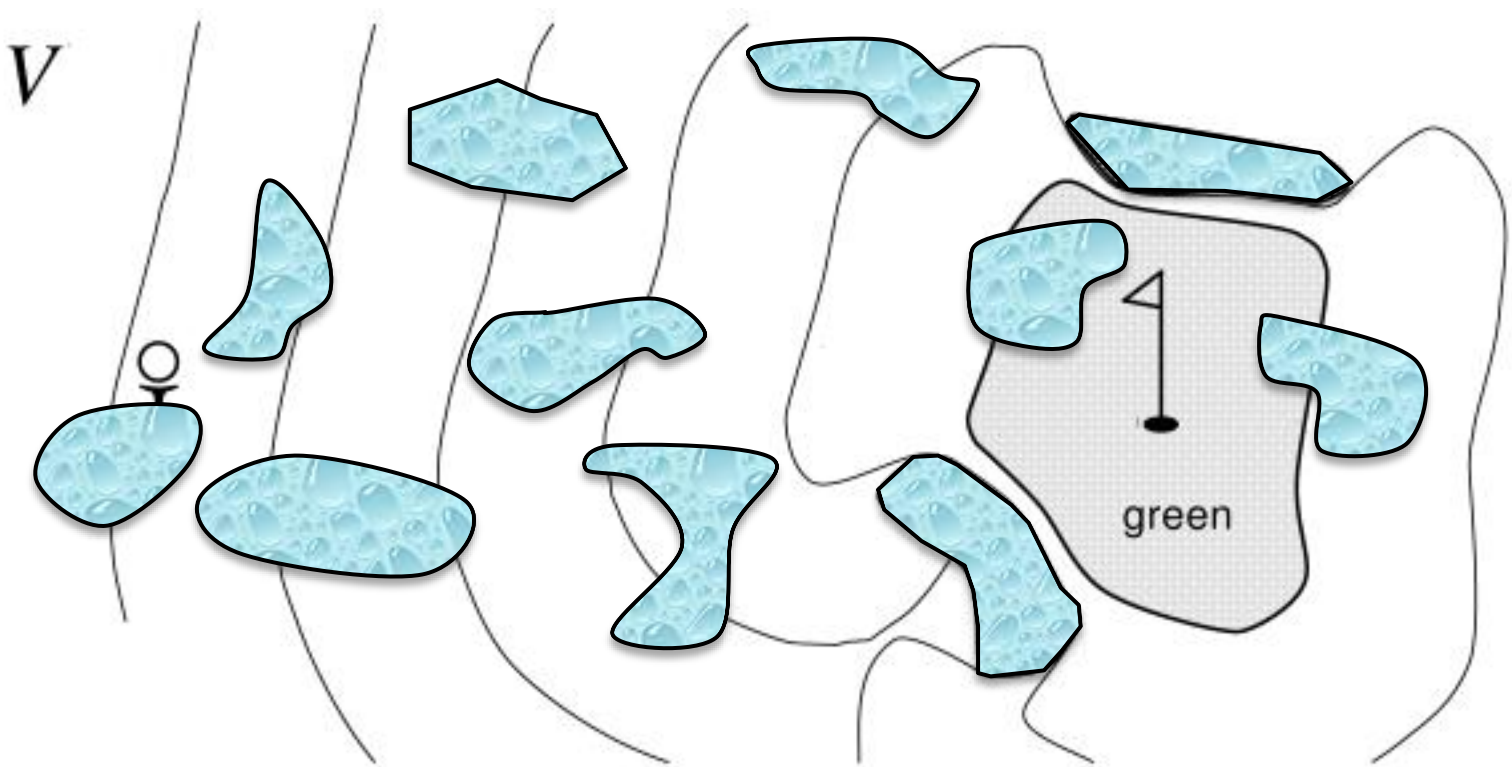
其中 $x=(x_1, x_2, \ldots, x_n)^T \in R^n$,

$f: R^n \rightarrow R^1$,

$g_i: R^n \rightarrow R^1$, $i=1, \ldots, p$,

$h_j: R^n \rightarrow R^1$, $j=1, \ldots, q$.

约束优化问题的困难



可行解?

可行方向?

解决方案

- 限定条件

  线性约束、凸规划……

- 问题转化

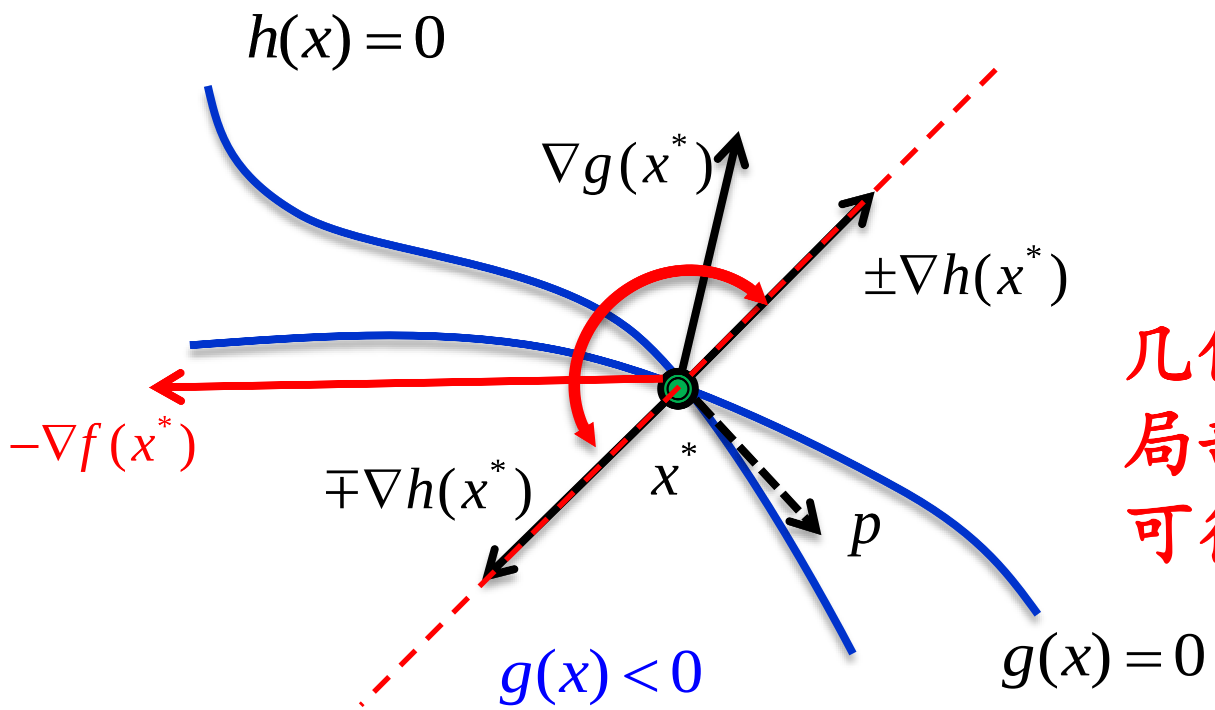
  约束优化->无约束优化

常用约束优化方法简介

- 罚函数类方法
  - 外点法, 内点法, 混合法, 精确罚函数法, 乘子(罚函数)法
  - 特点: 不能保证解的可行性. 乘子法效果最好
- 序列二次规划
  - (约束变尺度方法)
  - 特点: 不能保证解的可行性. 对中小问题效果很好
- 可行方向法
  - 投影梯度法, 简约梯度法(均要求线性约束)
  - 特点: 能保证解的可行性, 效果很好
- 特殊结构算法
  - 二次规划, 几何规划……

约束优化问题最优点的几何意义

$$\min \quad z = f(x)$$
$$s.t. \quad g(x) \le 0$$
$$h(x) = 0$$



几何意义：
局部最优解处不存在可行的下降方向！

如果$\nabla g(x^*)$，$\nabla h(x^*)$线性无关

则存在$\lambda \ge 0, \mu \in R$使得$-\nabla f(x^*) = \lambda \nabla g(x^*) + \mu \nabla h(x^*)$

基本概念

$$(NLP)\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\le 0;\quad i=1,2,\cdots,p \\ \qquad h_j(x)=0;\quad j=1,2,\cdots,q \\ \quad x\in R^n\end{cases}$$

$I=\{i\mid i=1,2,\cdots,p\}$ 不等式约束下标集

$E=\{j\mid j=1,2,\cdots,q\}$ 等式约束下标集

$D=\{x\mid x\in R^n,\ g_i(x)\le 0,\ h_j(x)=0,\ \forall i\in I,\ j\in E\}$ 可行域

$\forall x\in D,\ I(x)=\{i\mid i\in I,\ g_i(x)=0\}$ 有效不等式约束下标集

$L(x,\lambda,\mu)=f(x)+\sum_{i=1}^{p}\lambda_i g_i(x)+\sum_{j=1}^{q}\mu_j h_j(x)$ **Lagrange函数**

$\lambda=\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_p\end{pmatrix}^T\in R^p,\ \lambda\ge 0$ 不等式约束乘子(要求非负)

$\mu=\begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_q\end{pmatrix}^T\in R^q$ 等式约束乘子

定理1: K-T条件

$$\min \quad z = f(x)$$

- 对于非线性规划问题:
$$s.t. \quad g_i(x) \le 0$$
$$h_j(x) = 0$$
- 假设:
  - 1. f, g_i, h_j都是连续可微函数
  - 2. x^*是该问题的一个局部极小点
  - 3. $\{\nabla g_i(x^*)\}_{i \in I(x^*)}$, $\{\nabla h_j(x^*)\}_{j \in E}$线性无关
- 则存在一组$\lambda_i(i \in I(x^*))$, $\lambda_i \ge 0$和一组$\mu_j(j \in E)$, 使得

$$\nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{q} \mu_j \nabla h_j(x^*) = 0$$

几点说明

- 仅是x^*成为局部最优解的必要条件
- (3)中的条件称为约束规范
- 若令$i \in I\backslash I(x^*)$ 时$\lambda_i=0$, 则结论可写为:

$$\nabla f(x^*)+\sum_{i=1}^{p}\lambda_i\nabla g_i(x^*)+\sum_{j=1}^{q}\mu_j\nabla h_j(x^*)=0$$

或

$$\nabla_x L(x^*,\lambda,\mu)=0,\lambda_i g_i(x^*)=0$$

- 条件$\lambda_i g_i(x^*)=0$称为互补松弛条件
- KT点: $x\in R^n$, 满足:

$$\begin{cases}\nabla_x L(x,\lambda,\mu)=0\\ \lambda_i g_i(x)=0\end{cases}$$

定理2

定理2. 假设(NLP)中f, g_i (i=1,2,…,p)是可微凸函数，h_j (j=1,2,…,q)是线性函数，x^*是一个可行的KT点，则x^*是该问题的全局最优解。

证明：x^*为KT点 $\Rightarrow$ 存在$\lambda^* \geq 0, \mu^* \in R^q$使得$\nabla_x L(x^*, \lambda^*, \mu^*) = 0$

$$L(x,\lambda^*,\mu^*) = f(x) + \sum_{i=1}^{p} \lambda_i^* g_i(x) + \sum_{j=1}^{q} \mu_j^* h_j(x)\text{关于}x\text{为凸函数}$$

$$f(x^*) = L(x^*,\lambda^*,\mu^*) \leq L(x,\lambda^*,\mu^*) \leq f(x), \quad \forall x \in D$$

可行KT点，互补松弛条件

x^*为L的全局极小

x可行，乘子非负

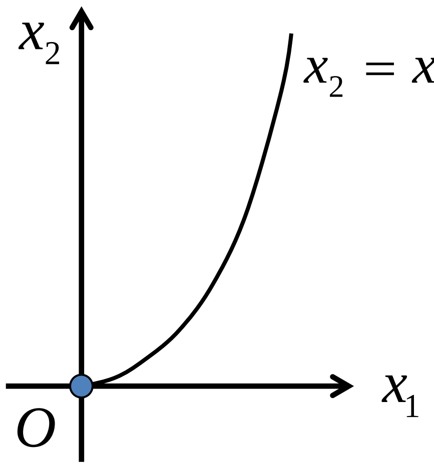
定理2

- 对于一个NLP问题,若$f(x)$, $g_i(x)$, $i=(1,\ldots, p)$都是凸函数, $h_j(x)$, $j=(1,\ldots, q)$都是线性函数. 且x^*是一个可行KT点, 则x^*是该问题的全局最优解

例1

$$\begin{aligned} \min \quad & x_1 \\ s.t. \quad & x_2 \le x_1^3 \\ & x_2 \ge 0 \end{aligned}$$

例1：$\begin{cases} \min & x_1 \\ s.t. & x_2 \le x_1^3 \\ & x_2 \ge 0 \end{cases} \Rightarrow \begin{cases} \min & x_1 \\ s.t. & x_2 - x_1^3 \le 0 \\ & -x_2 \le 0 \end{cases}$



$$x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla f(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla g_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

线性相关，约束规范不成立！

不存在$\lambda_1, \lambda_2 \ge 0$使得$\nabla f(x^*) + \lambda_1 \nabla g_1(x^*) + \lambda_2 \nabla g_2(x^*) = 0$

约束规范不成立时，局部最优解可能不是KT点！

例2 用K-T条件求解问题

$$
\begin{aligned}
\min \quad & f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2 \\
s.t. \quad & g_1(x) = x_1 + x_2 - 2 \le 0 \\
& g_2(x) = -x_1 \le 0 \\
& g_3(x) = -x_2 \le 0 \\
& h_1(x) = -x_1 + x_2 - 1 = 0
\end{aligned}
$$

例2:

解: 先写出Lagrange函数.

$$L(x,\lambda,\mu)=(x_1-1)^2+(x_2-2)^2+\lambda_1(x_1+x_2-2)-\lambda_2 x_1-\lambda_3 x_2+\mu(-x_1+x_2-1)$$

依次写出KT条件、乘子非负条件、可行条件(原问题约束)

$$\begin{cases}\begin{cases}2(x_1-1)+\lambda_1-\lambda_2-\mu=0\\ 2(x_2-2)+\lambda_1-\lambda_3+\mu=0\end{cases}\\ \begin{cases}\lambda_1(x_1+x_2-2)=0\\ \lambda_2 x_1=0\\ \lambda_3 x_2=0\end{cases}\\ \{\ \lambda_1,\lambda_2,\lambda_3\ge 0\\ \begin{cases}x_1+x_2-2\le 0\\ -x_1+x_2-1=0\\ x_1,x_2\ge 0\end{cases}\end{cases}$$

分情况讨论：①. $\lambda_1\neq 0$

$$\begin{cases}\lambda_1(x_1+x_2-2)=0\\ -x_1+x_2-1=0\end{cases}\Rightarrow\begin{cases}x_1=1/2\\ x_2=3/2\end{cases}\Rightarrow\begin{cases}\lambda_2=0\\ \lambda_3=0\end{cases}$$

$$\begin{cases}\lambda_1=1\\ \mu=0\end{cases}$$ **已得KT点.** √

分情况讨论：②. $\lambda_1=0$

$$\begin{cases}x_1,x_2\ge 0\\ -x_1+x_2-1=0\end{cases}\Rightarrow\begin{cases}\text{或者}x_1>0,x_2>0\\ \text{或者}x_1=0,x_2=1\end{cases}$$

$$\Rightarrow\begin{cases}\text{或者}x_1=1,x_2=2,\lambda_2=\lambda_3=0,\mu=0\\ \text{或者}x_1=0,x_2=1,\lambda_3=0,\lambda_2=-4,\mu=2\end{cases}$$ ×

作业

- P155 21, 22(1)