对偶理论（Duality）

张伯雷

南京邮电大学 计算机学院、通达学院

https://bolei-zhang.github.io/course/opt.html

回顾

- 凸集
  - 集合中任意两点所组成的线段仍然在该集合中
- 凸函数
  - 定义一：如果dom$f$为凸集，且$f(\theta x+(1-\theta)y)\le\theta f(x)+(1-\theta)f(y)$对所有的$x,y\in\mathrm{dom}f$，$0\le\theta\le 1$成立
  - 定义二：对于可微函数f，如果dom$f$为凸集，则f为凸函数当且仅当$f(y)\ge f(x)+\nabla f(x)^T(y-x)$，对所有的$x,y\in\mathrm{dom}f$成立
  - 定义三：对于二阶可微函数f，如果dom$f$为凸集，则f为凸函数当且仅当$\nabla^2 f(x)\succcurlyeq 0$，对所有$x\in\mathrm{dom}f$成立
- 凸优化问题
  - 标准形式
  - 局部最优=全局最优
  - 最优条件：$x\in X$最优 $\Leftrightarrow \nabla f_0^T(x)(y-x)\ge 0$ for all $y\in X$

目录

- 拉格朗日与共轭函数
- 对偶间隙
- 最优性条件：Slater条件
- 最优条件的两个解释
  - 几何解释
  - 鞍点解释
- 互补松弛与KKT条件

一般优化问题

- 优化问题

$$\min f_0(x).$$
$$\text{s.t.}\quad f_i(x) \le 0, \qquad i = 1, \ldots, m$$
$$h_i(x) = 0, \qquad i = 1, \ldots, p$$

  - $x \in R^n$
  - $D = \bigcap_{i=0}^{m} dom f_i \cap \bigcap_{i=1}^{p} dom h_i$
  - $X = \{x \in D$ 且所有约束条件可以满足$\}$
  - p^* 为最优值

不一定是凸优化问题!

拉格朗日函数

- 优化问题

$$\begin{aligned} \min \quad & f_0(x). \\ \text{s.t.} \quad & f_i(x) \le 0, \qquad i = 1, \ldots, m \\ & h_i(x) = 0, \qquad i = 1, \ldots, p \end{aligned}$$

- 拉格朗日函数（Lagrangian function）
  - 给该问题中的每一个约束指定一个拉格朗日乘子λ, ν，以乘子为加权系数将约束增加到目标函数中
  - L: $\mathbb{R}^n \times \mathbb{R}^{\mathrm{m}} \times \mathbb{R}^p \to \mathbb{R}$，其中定义域为$D \times \mathbb{R}^{\mathrm{m}} \times \mathbb{R}^p$，

$$L(x, \lambda, \upsilon) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \upsilon_i h_i(x)$$

拉格朗日对偶函数

- 拉格朗日对偶函数(Lagrange Function/Dual Function)

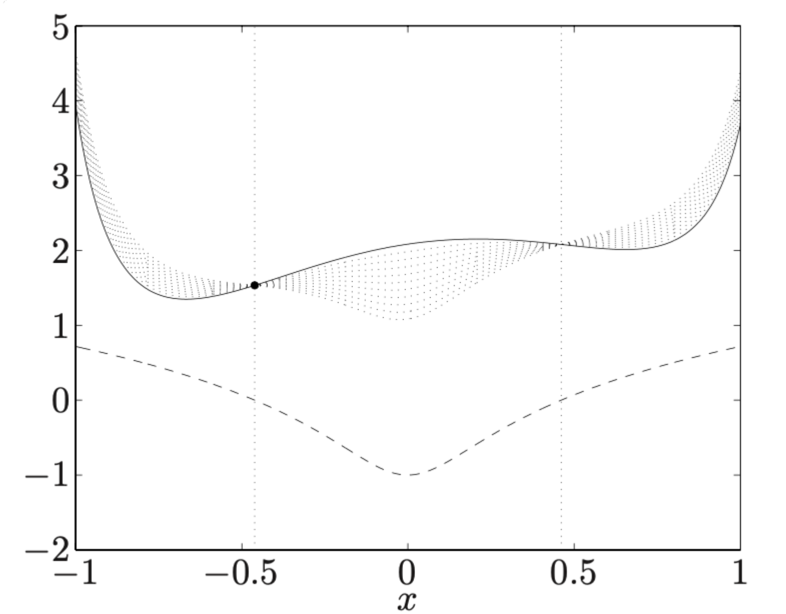
$$g(\lambda, v) = \inf_{x \in D} L(x, \lambda, v)$$

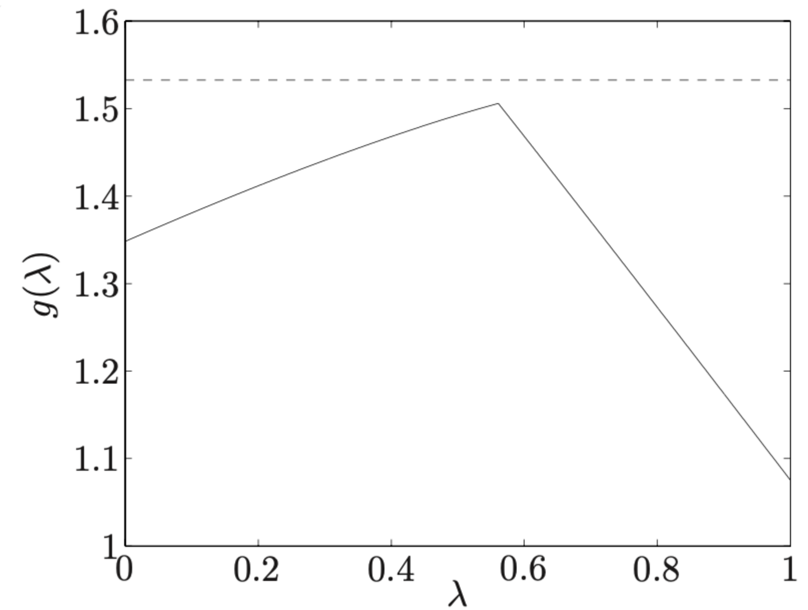
$$= \inf_{x \in D} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)\right)$$

- $\lambda \in \mathbb{R}^m, v \in \mathbb{R}^p$

对偶函数性质一

- 弱对偶理论：如果$\lambda \geq 0$，则$g(\lambda, \upsilon) \leq p^*$





实线：目标函数f_0

虚线：$f_1 \leq 0$

点虚线：拉格朗日$f_0 + \lambda f_1$:$\lambda = 0.1, 0.2, \ldots, 1.0$

可行解集：[-0.46, 0.46]

对偶函数性质二

- 性质：对偶函数为凹函数

$$g(\lambda,\nu)=\inf_{x\in\mathcal{D}} L(x,\lambda,\nu)=\inf_{x\in\mathcal{D}}\left(f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p}\nu_i h_i(x)\right)$$

- 命题：若$f_1,\dots,f_m$为凸函数，则$f(x)=\max\{f_1(x),\dots,f_m(x)\}$为凸函数
- 命题：若$f(x,y)$对所有$y\in A$为关于x的凸函数，则$f(x)=\sup_{y\in A} f(x,y)$ 为凸函数

例1

$$\min \quad x^T x$$

$$\text{s.t.} \quad \mathrm{Ax} = b$$

$$x \in \mathbb{R}^n, b \in \mathbb{R}^p, A \in \mathbb{R}^{p \times n}$$

例2

$$\min \quad C^T x$$

$$\text{s.t.} \quad Ax = b\,, x \geq 0$$

例3

$$\begin{array}{ll} \min & x^T W x \\ \text{s.t.} & x_i = \pm 1, i = 1, \dots, m \end{array}$$

函数的共轭

- 函数共轭：

$$f^*\text{是}f\text{的共轭，若}f^*(y)=\sup_{x\in dom\, f}(y^Tx-f(x))$$

- 拉格朗日对偶函数(Lagrange Function/Dual Function)

$$g(\lambda,v)=\inf_{x\in D}L(x,\lambda,v)$$

$$=\inf_{x\in D}(f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p}v_i h_i(x))$$

例

$$\begin{aligned} \min \quad & f(x) \\ \mathrm{s.t.} \quad & x = 0 \end{aligned}$$

- 解

例

$$\begin{aligned} \min \quad & f_0(x) \\ s.t. \quad & Ax \le b, Cx = d \end{aligned}$$

- 解