第十三讲：对偶

优化问题的对偶性

杨 林

大 纲

1. 拉格朗日对偶问题

2. 弱对偶和强对偶

3. 例子:强对偶与弱对偶

大 纲

1. 拉格朗日对偶问题

2. 弱对偶和强对偶

3. 例子:强对偶与弱对偶

1 拉格朗日对偶问题

■ **拉格朗日函数：给定标准形式问题(不一定是凸的)**

$$
\begin{array}{ll}
\text{最小化} & f_0(x) \\
\text{约束条件} & f_i(x) \leq 0,\ i = 1, 2, \dots, m \\
& h_i(x) = 0,\ \ i = 1, 2, \dots, p
\end{array}
$$

变量 $x \in \mathbb{R}^n$，定义域 $\mathcal{D}$，最优值 p^*

1 拉格朗日对偶问题

■ 拉格朗日函数（Lagrange）：

L： $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$，其中 $\mathbf{dom}\, L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

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

- 目标函数和约束函数的加权和
- λ_i 是与 $f_i(x) \leq 0$ 相关的拉格朗日乘子
- ν_i 是与 $h_i(x) = 0$ 相关的拉格朗日乘子

1 拉格朗日对偶问题

■ **拉格朗日对偶函数**： $g\colon\ \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

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

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

■ g 是凹函数

■ **下界性质**：如果 $\lambda \geq 0$, 那么 $g(\lambda, \nu) \leq p^*$

□ **证明**：如果 $\tilde{x}$ 是可行的且 $\lambda \geq 0$, 那么

$$f_0(\tilde{x}) \geq L(\tilde{x}, \lambda, \nu) \geq \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = g(\lambda, \nu)$$

对所有的可行的 $\tilde{x}$ 进行最小化得到 $p^* \geq g(\lambda, \nu)$

1 拉格朗日对偶问题

■ 线性方程的最小范数解

最小化	$x^T x$
约束条件	$Ax = b$

□ 对偶方程:

1. 拉格朗日函数是$L(x,\nu) = x^T x + \nu^T(Ax - b)$
2. 要在 x 上最小化 L， 设梯度等于零:

$$\nabla_x L(x,\nu) = 2x + A^T\nu = 0 \Longrightarrow x = -(1/2)A^T\nu$$

代入 L 中得到 g :

$$g(\nu) = L\big((-1/2)A^T\nu, \nu\big) = -(1/4)\nu^T AA^T\nu - b^T\nu$$

ν 的凹函数

□ 下界性质: $p^* \geq -(1/4)\nu^T AA^T\nu - b^T\nu$ 对于所有 ν

1 拉格朗日对偶问题

■ 标准形式 LP

最小化	$c^T x$
约束条件	$Ax = b$, $x \geq 0$

1. 拉格朗日函数是

$$\begin{aligned} L(x,\lambda,\nu) &= c^T x + \nu^T (Ax - b) - \lambda^T x \\ &= -b^T \nu + (c + A^T \nu - \lambda)^T x \end{aligned}$$

2. L 是 x 的仿射函数，因此

$$g(\lambda,\nu) = L(x,\lambda,\nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{其他情况} \end{cases}$$

g 在仿射域 $\{(\lambda,\nu) | A^T \nu - \lambda + c = 0\}$ 上是线性的, 因此是凹的

□ 下界性质： $p^* \geq -b^T \nu$ 对于所有 ν 若 $A^T \nu + c \geq 0$

1 拉格朗日对偶问题

■ 等式约束范数最小化

最小化	$\|x\|$
约束条件	$Ax = b$

□ 对偶方程:

$$g(\nu) = \inf_x(\| x \| - \nu^T Ax + b^T \nu) = \begin{cases} b^T \nu & \|A^T \nu\|_* \leq 1 \\ -\infty & \text{其他情况} \end{cases}$$

其中 $\|\nu\|_* = \sup\limits_{\|u\| \leq 1} u^T \nu$ 是 $\|\cdot\|$ 的对偶范数

1 拉格朗日对偶问题

■ 等式约束范数最小化(续)

□ 证明:

当 $\|y\|_* \leq 1$ 时, $\inf_x(\|x\| - y^T x) = 0$;否则为 $-\infty$

1. 如果 $\|y\|_* \leq 1$, $\inf_x(\|x\| - y^T x) = \inf_x(1 - y^T x/\|x\|)\,\|x\|$ ≥ 0 , 当 $x = 0$ 时取等号

2. 如果 $\|y\|_* > 1$, 令 $x = tu$, 其中$\|u\| \leq 1$, $u^T y = \|y\|_* > 1$:

$$\|x\| - y^T x = t(\|u\| - \|y\|_*) \to -\infty \text{ 随着 } t \to \infty$$

□ 下界性质: $p_* \geq b^T \nu$ 若 $\|A^T \nu\|_* \leq 1$

1 拉格朗日对偶问题

■ 两向划分

最小化	$x^T W x$
约束条件	$x_i^2 = 1, i = 1, \cdots n$

1. 非凸问题；可行集包含2^n个离散点
2. 解释：将$\{1, \cdots, n\}$划分为两个集合；W_{ij} 是 i 和 j 分配到同一集合的成本；$-W_{ij}$ 是分配到不同集合的成本

□ 对偶方程：

$$g(\nu) = \inf_x \left((x^T W x + \sum_i \nu_i (x_i^2 - 1)) \right)$$
$$= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu$$

1 拉格朗日对偶问题

■ 两向划分(续)

□ **对偶方程:**

$$g(\nu)=\begin{cases}1^T\nu & W+\mathbf{diag}(\nu)\succeq 0\\ -\infty & \text{其他情况}\end{cases}$$

□ **下界性质:** $p^*\ge -\mathbf{1}^T\nu$ 若 $W+\mathbf{diag}(\nu)\succeq 0$

□ **例:** $\nu=-\lambda_{\min}(W)\mathbf{1}$ 给出界限 $p^*\ge n\lambda_{\min}(W)$

1 拉格朗日对偶问题

■ 拉格朗日对偶与共轭函数

最小化	$f_0(x)$
约束条件	$Ax \preceq b,\ Cx = d$

□ 对偶方程:

$$g(\nu) = \inf_{x \in \mathbf{dom}\, f_0} \left(f_0(x) + (A^T\lambda + C^T\nu)^T x - b^T\lambda - d^T\nu\right)$$

$$= -f_0^*(-A^T\lambda - C^T\nu) - b^T\lambda - d^T\nu$$

1. 回顾共轭的定义 $f^*(y) = \inf_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$
2. 若已知 f_0 的共轭, 可简化对偶的推导

1 拉格朗日对偶问题

■ 对偶问题

最小化	$g(\lambda,\nu)$
约束条件	$\lambda \succeq 0$

1. 找到 p^* 的最佳下界, 该下界来自拉格朗日对偶函数
2. 一个凸优化问题； 最优值用 d^* 表示
3. λ,ν 是对偶可行的, 如果 $\lambda \succeq 0$, $(\lambda,\nu) \in \mathbf{dom}\ g$
4. 通常通过将隐式约束 $(\lambda,\nu) \in \mathbf{dom}\ g$ 显式化来简化问题

□ 例 1 : 标准形式 LP 及其对偶

最小化	c^Tx	最大化	$-b^Tx$
约束条件	$Ax = b,$	约束条件	$A^T\nu + c \succeq 0$
	$x \succeq 0$		

大 纲

1. 拉格朗日对偶问题

2. 弱对偶和强对偶

3. 例子:强对偶与弱对偶

2 弱对偶和强对偶

■ **弱对偶：** $d^* \leq p^*$

1. 始终成立（对凸和非凸问题）
2. 可用于为困难问题找到非平凡的下界，例如求解

$$\begin{array}{ll} \text{最大化} & -\mathbf{1}^T v \\ \text{约束条件} & W + \mathbf{diag}(v) \succeq 0 \end{array}$$

给出双向划分问题的下界

■ **强对偶：** $d^* = p^*$

1. 一般不成立
2. （通常）适用于凸问题
3. 保证凸问题强对偶性的条件被称为**约束准则**

2 弱对偶和强对偶

■ **Slater约束准则**

对于以下凸问题

$$
\begin{array}{ll}
\text{最小化} & f_0(x) \\
\text{约束条件} & f_i(x) \le 0, \quad i = 1, \cdots, m \\
 & Ax \ = \ b
\end{array}
$$

如果它是严格可行的, 即

$$\exists x \in \mathbf{int}\,\mathcal{D}: \quad f_i(x) < 0, i = 1, \cdots, m, Ax = b$$

强对偶性成立

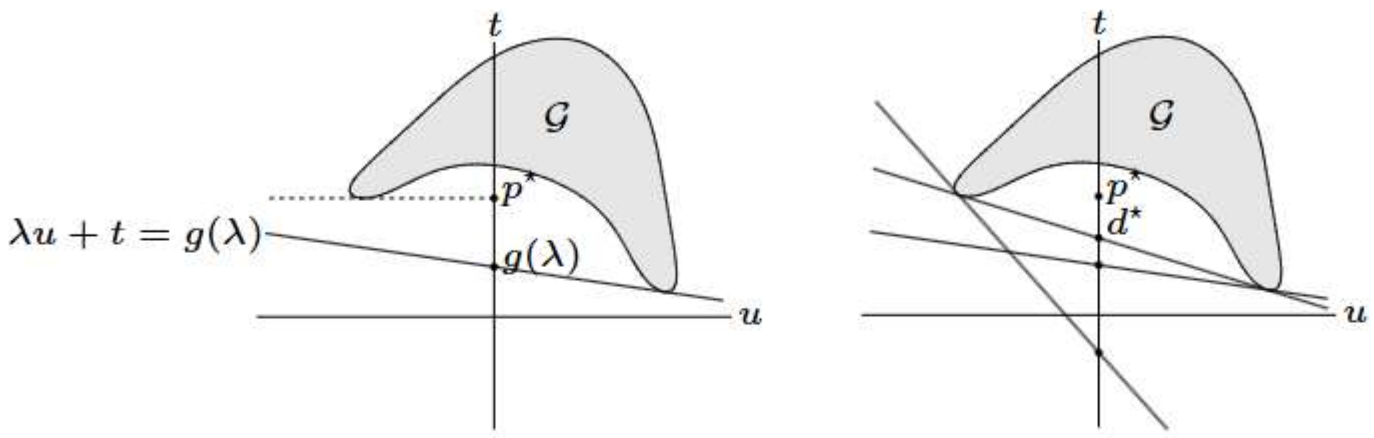
■ 存在许多其他类型的约束条件

2 弱对偶和强对偶

为简化起见，考虑具有一个约束条件 $f_1(x) \le 0$ 的问题

■ **对偶函数的几何解释:**

$$g(\lambda) = \inf_{(u,t)\in\mathcal{G}} (t + \lambda u)\,,\mathcal{G} = \left\{\left(f_1(x),\ \ f_0(x)\right)\middle| x \in \mathcal{D}\right\}$$

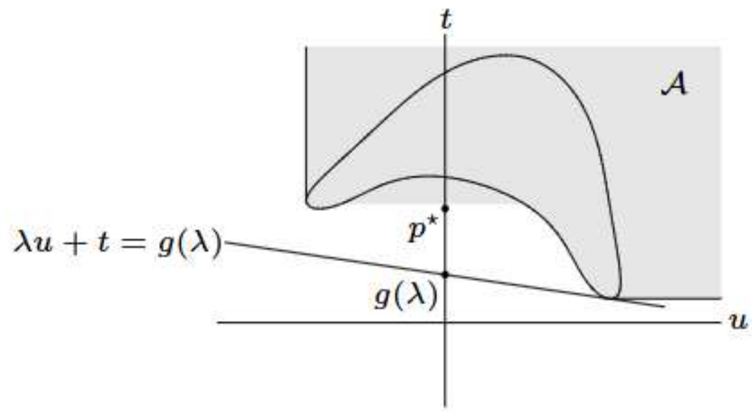


1. $\lambda u + t = g(\lambda)$ 是 $\mathcal{D}$ 的（非垂直）支持超平面
2. 超平面与 t 轴相交于 $t = g(\lambda)$

2 弱对偶和强对偶

- **上镜图解释：** 如果用上镜图$\mathcal{A}$替换 $\mathcal{G}$，解释相同

$$\mathcal{A} = \{(u,t) | f_1(x) \le u, f_0(x) \le t \text{ 对于某些 } x \in \mathcal{D}\}$$



- **强对偶性:**

1. 当 $(0, p^*)$ 处存在 $\mathcal{A}$ 的非垂直支撑超平面时强对偶性才会成立
2. 对于凸问题，$\mathcal{A}$ 是凸的，因此在 $(0, p^*)$ 处有支撑超平面
3. 斯莱特条件：如果 $(\tilde{u}, \tilde{t}) \in \mathcal{A}$ 其中 $\tilde{u} < 0$，那么在 $(0, p^*)$ 处的支持超平面必须是非垂直的

2 弱对偶和强对偶

- **数学证明:**

1. 包含非空内部点;

2. rank $A = p$. 定义以下两个互斥的集合:

$$\mathcal{A} = \{(u, v, t) | \exists x \in \mathcal{D}, f_i(x) \leq u_i, h_i(x) = v_i, f_0(x) \leq tx\}$$

$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^*\}$$

$\exists(\tilde{\lambda}, \tilde{v}, \mu) \neq 0$ 和 α, 使得

$$(u, v, t) \in \mathcal{A} \Longrightarrow \tilde{\lambda}^T u + \tilde{v}^T v + \mu t \geq \alpha$$

$$(u, v, t) \in \mathcal{B} \Longrightarrow \tilde{\lambda}^T u + \tilde{v}^T v + \mu t \leq \alpha$$

我们有: $1. \tilde{\lambda} \geq 0$ 以及 $\mu \geq 0$ (由于上图的特性); $2. \mu p^* \leq \alpha$

换句话说, 存在可行解 $x \in \mathcal{D}$

$$\sum \tilde{\lambda}_i f_i(x) + \tilde{v}^T (Ax - b) + \mu f_0(x) \geq \alpha \geq \mu p^*$$

2 弱对偶和强对偶

情况 1: $\mu > 0$

$L(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^*$, 因此 $g(\tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^*$

$g(\tilde{\lambda}/\mu, \tilde{\nu}/\mu) = p^*$, 证毕

情况 2: $\mu = 0$

$$\sum \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T(Ax - b) \geq 0, \forall x \in \mathcal{D}$$

假设 $\tilde{x}$ 满足斯莱特条件. 那么

$$A\tilde{x} - b = 0, \sum \tilde{\lambda}_i f_i(x) \geq 0 \Longrightarrow \tilde{\lambda} = 0$$

2 弱对偶和强对偶

因此, $\tilde{v} \neq 0$ (超平面的非零参数)

考虑以下事实:

1. $\tilde{v}(A\tilde{x} - b) = 0$
2. $\tilde{x} \in \mathbf{int}\,\mathcal{D}$

除非 $\mathrm{A}^{\mathrm{T}}\tilde{v} = 0$, 否则存在 x, 使得 $\tilde{v}(Ax - b) < 0$ $(\mathbf{rank}\,A = p, \tilde{v}$ 是 p 维的)

大 纲

1. 拉格朗日对偶问题

2. 弱对偶和强对偶

3. 例子:强对偶与弱对偶

3 例子：强对偶与弱对偶

- 例 2:不等式形式 ***LP***

- **原问题:**

$$\begin{array}{ll} \text{最小化} & c^T x \\ \text{约束条件} & Ax \le b \end{array}$$

- **对偶函数:**

$$g(\lambda) = \inf_x \left((c + A^T\lambda)^T x - b^T\lambda \right) = \begin{cases} -b^T\lambda & A^T\lambda + c = 0 \\ -\infty & \text{其他情况} \end{cases}$$

3 例子：强对偶与弱对偶

- **例 2:不等式形式 *LP***

- **对偶问题:**

$$\begin{array}{ll} \text{最小化} & -b^T\lambda \\ \text{约束条件} & A^T\lambda + c = 0, \lambda \geq 0 \end{array}$$

1. 根据斯莱特条件: $p^* = d^*$,如果 $A\tilde{x} \prec b$ 对于某些 $\tilde{x}$
2. 实际上, $p^* = d^*$ 除非原始问题和对偶问题都是不可行的

4 例子：强对偶与弱对偶

- 例 3:二次规划

- **原问题:**

最小化 $x^T P x$
约束条件 $Ax \preceq b$

- **对偶函数:**

$$g(\lambda) = \inf_x \left(x^T P x + \lambda^T (AX - B) \right) = -\frac{1}{4} \lambda^T A P^{-1} A^T \lambda - b^T \lambda$$

3 例子：强对偶与弱对偶

- □ **例 3:二次规划**
- ■ **对偶问题:**

$$\begin{array}{ll} \text{最小化} & -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{约束条件} & \lambda \succeq 0 \end{array}$$

1. 根据斯莱特条件: $p^* = d^*$,如果 $A\tilde{x} \prec b$ 对于某些 $\tilde{x}$
2. 实际上, $p^* = d^*$ 总是成立

3 例子：强对偶与弱对偶

- 例 4:一个具有强对偶性的非凸问题:

  - **原问题:**

$$
\begin{array}{ll}
\text{最小化} & x^T A x + 2b^T x \\
\text{约束条件} & x^T x \leq 1
\end{array}
$$

$A \not\succeq 0$, 所以非凸

  - **对偶函数:**

$$g(\lambda) = \inf_x (x^T (A + \lambda I) x + 2b^T x - \lambda)$$

1. 若 $A + \lambda I \not\succeq 0$ 或若 $A + \lambda I \succeq 0$ 且 $b \notin \mathcal{R}(A + \lambda I)$, 则无界下界(以下无界)
2. 由 $x = -(A + \lambda I)^\dagger b$ 最小化, 否则: $g(\lambda) = -b^T (A + \lambda I)^\dagger b - \lambda$

3 例子：强对偶与弱对偶

- **例 4:一个具有强对偶性的非凸问题:**

- **对偶问题与等价 SDP:**

$$
\begin{array}{ll}
\text{最小化} & -b^T(A+\lambda I)^{\dagger}b-\lambda \\
\text{满足} & A+\lambda I \succeq 0 \\
 & b \in \mathcal{R}(A+\lambda I)
\end{array}
$$

$$
\begin{array}{ll}
\text{最大化} & -t-\lambda \\
\text{满足} & \begin{bmatrix} A+\lambda I & b \\ b^T & t \end{bmatrix}
\end{array}
$$

强对偶性, 尽管原始问题不是凸的(不容易证明)

谢 谢 ！