一般的约束优化问题

$$\min_{x} f(x)$$

index set　　m个约束

$$\text{s.t.}\quad c_1(x) \le 0 \Rightarrow c_i(x) \le 0 \quad i \in \mathcal{I}, \quad |\mathcal{I}| = m$$

$$c_2(x) = 0 \Rightarrow c_j(x) = 0 \quad j \in \mathcal{E}, \quad |\mathcal{E}| = p$$

最优值：$z^* = \min_{x \in \Omega} f(x)$

可行域 $\Omega = \{x \in \mathbb{R}^n \mid c_1(x) \le 0,\ c_2(x) = 0\}$　　$\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$

拉格朗日函数：$L(x; \lambda, \mu) = f(x) + \lambda^T c_1(x) + \mu^T c_2(x)$

$$= f(x) + \sum_{i \in \mathcal{I}} \lambda_i c_{1i}(x) + \sum_{j \in \mathcal{E}} \mu_i c_{2i}(x)$$

拉格朗日对偶函数：

$$g(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x; \lambda, \mu)$$

$$= \inf_{x \in \mathbb{R}^n} \left(f(x) + \lambda^T c_1(x) + \mu^T c_2(x) \right)$$

E.g.

$$\begin{aligned} \max\ & c^T x \\ \text{s.t.}\ & Ax = b \\ & x \ge 0 \end{aligned} \quad\Rightarrow\quad \begin{aligned} \min\ & b^T y \\ \text{s.t}\ & A^T y \ge c \\ & y \ge 0 \end{aligned}$$

$$\Rightarrow \quad \begin{aligned} \min\ & -c^T x \\ \text{s.t}\ & Ax = b \\ & -x \le 0 \end{aligned}$$

拉格朗日函数：

$$L(x; \lambda, \mu) = -c^T x + \lambda^T(-x) + \mu^T(Ax - b)$$

$$= (A^T\mu - \lambda - c)^T x - b^T \mu$$

拉格朗日对偶函数：

$$g(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x; \lambda, \mu) = \begin{cases} -b^T\mu & \text{if } A^T\mu - \lambda - c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

case 1　$A^T\mu - \lambda - c = 0$　$L = -b^T\mu$

case 2　$A^T\mu - \lambda - c \ne 0$　$L = (A^T\mu - \lambda - c)^T x - b^T\mu$

if $A^T\mu - \lambda - c > 0$　$x \to -\infty$　$L \to -\infty$

if $A^T\mu - \lambda - c < 0$　$x \to \infty$　$L \to -\infty$

$\lambda \geq 0$ $\quad$ $\lambda = A^T\mu - c \geq 0$

拉格朗日对偶问题:

$$\max \; -b^T\mu \qquad \text{s.t.} \; A^T\mu - c - \lambda = 0, \; \lambda \geq 0$$

$$\max \; -b^T\mu \qquad \text{s.t.} \; A^T\mu - c \geq 0$$

$$\Rightarrow \min \; b^T\mu \qquad \text{s.t.} \; A^T\mu \geq c$$

Thm (Weak duality Theorem) 若 $\lambda \geq 0$, 则 $g(\lambda, \mu) \leq \underline{z^*}$ → 原问题最优值

Proof: 若 $\bar{x} \in \Omega$, 则 $c_1(\bar{x}) \leq 0$, $c_2(\bar{x}) = 0$

$$L(\bar{x}; \lambda, \mu) = f(\bar{x}) + \lambda^T \underbrace{c_1(\bar{x})}_{\leq 0} + \mu^T \underbrace{c_2(\bar{x})}_{=0} \leq f(\bar{x})$$

而 $g(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x; \lambda, \mu) \leq L(\bar{x}; \lambda, \mu) \leq f(\bar{x})$

对 $\bar{x}$ 取下界 $\quad f(\bar{x}) = \min_{x \in \Omega} f(x) = z^*$

$$\Rightarrow g(\lambda, \mu) \leq z^*$$

拉格朗日对偶问题

$$\max_{\substack{\lambda \geq 0 \\ \mu \in \mathbb{R}^p}} g(\lambda, \mu) = \max_{\substack{\lambda \geq 0 \\ \mu \in \mathbb{R}^p}} \inf_{x \in \mathbb{R}^n} L(x; \lambda, \mu)$$

对偶变量: λ, μ

最优值: w^* → w^* 为 z^* 的最优下界, $z^* - w^*$ 对偶间隙 dual gap

Prop ① Weak duality inequality $z^* \geq w^*$ 一定成立

② 对偶问题一定是一个凸问题 (= easy to solve)

→ 目标函数为凸 (最小化问题) 且可行域为凸集

对偶问题可以为原问题提供一个有用的下界.

E.x. 求下列问题的拉格朗日对偶问题:

$$\begin{aligned} \max\ & c^T x \\ s.t\ & A_1 x = b_1 \\ & A_2 x \le b_2 \\ & A_3 x \ge b_3 \\ & x \ge 0 \end{aligned} \Rightarrow \begin{aligned} \min\ & -c^T x \\ s.t\ & \begin{array}{cc} \begin{pmatrix} A_2 \\ -A_3 \\ -I \end{pmatrix} x \le \begin{pmatrix} b_2 \\ -b_3 \\ 0 \end{pmatrix} & \begin{matrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{matrix} \end{array} \\ & A_1 x = b_1 \qquad \mu \end{aligned}$$

$$A = \begin{pmatrix} A_2 \\ -A_3 \\ -I \end{pmatrix} \qquad b = \begin{pmatrix} b_2 \\ -b_3 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{aligned} \min\ & -c^T x \\ s.t\ & Ax \le b \\ & A_1 x = b_1 \end{aligned}$$

$$L(x; \lambda_1, \lambda_2, \lambda_3, \mu) = -c^T x + \lambda_1^T (A_2 x - b_2) - \lambda_2^T (A_3 x - b_3) - \lambda_3^T x + \mu^T (A_1 x - b_1)$$

$$g(\lambda_1, \lambda_2, \lambda_3, \mu) = \inf_x \left(A_2^T \lambda_1 - A_3^T \lambda_2 + A_1^T \mu - \lambda_3 - c\right)^T x - b_2^T \lambda_1 + b_3^T \lambda_2 - b_1^T \mu$$

$$= \begin{cases} -b_2^T \lambda_1 + b_3^T \lambda_2 - b_1^T \mu & \text{if } A_2^T \lambda_1 - A_3^T \lambda_2 + A_1^T \mu - \lambda_3 - c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\begin{aligned} \max\ & -b_2^T \lambda_1 + b_3^T \lambda_2 - b_1^T \mu \\ s.t.\ & A_2^T \lambda_1 - A_3^T \lambda_2 + A_1^T \mu - \lambda_3 - c = 0 \\ & \lambda_{1-3} \ge 0 \end{aligned}$$

$$\begin{aligned} \max\ & -b_2^T \lambda_1 + b_3^T \lambda_2 - b_1^T \mu \\ s.t.\ & A_2^T \lambda_1 - A_3^T \lambda_2 + A_1^T \mu \ge c \\ & \lambda_{1-2} \ge 0 \end{aligned}$$

E.g. $\min \|x\|_2$

s.t. $Ax=b$

$\mathcal{L}(x;\mu) = \|x\|_2 - \mu^T(Ax-b)$

$g(\mu) = \inf_{x\in\mathbb{R}^n} \left(\|x\|_2 - \mu^T(Ax-b)\right)$

$= \inf_{x\in\mathbb{R}^n} \left(\|x\|_2 - (A^T\mu)^T x\right) + b^T\mu = \begin{cases} b^T\mu & \text{if } \|A^T\mu\|_2 \le 1 \\ -\infty & \text{otherwise} \end{cases}$

$-\sup_{x\in\mathbb{R}^n} \left(y^Tx - \|x\|_2\right) \Rightarrow \|x\|_2 \leftrightarrow$ 共轭函数 $\quad y = A^T\mu$

$y^Tx - \|x\|_2 \le \|y\|_2\|x\|_2 - \|x\|_2 = (\|y\|_2 - 1)\|x\|_2$

① $\|y\|_2 > 1 \quad \sup_{x\in\mathbb{R}^n} (y^Tx - \|x\|_2) = +\infty$

② $\|y\|_2 \le 1 \quad \sup_{x\in\mathbb{R}^n} (y^Tx - \|x\|_2) = 0$ when $x=0$

$\max b^T\mu$

s.t. $\|A^T\mu\|_2 \le 1$

E.g. $\min_x x^TWx \qquad W \succ 0$

$Ax \le b$

$\mathcal{L}(x;\mu) = x^TWx + \mu^T(Ax-b) \qquad \mu \ge 0$

$g(\mu) = \inf_{x\in\mathbb{R}^n} \left(x^TWx + (A^T\mu)^T x\right) - b^T\mu$

$2Wx + A^T\mu = 0 \Rightarrow x = -\frac{1}{2}W^{-1}A^T\mu$

$g(\mu) = -\frac{1}{4}\mu^TAW^{-1}A^T\mu - b^T\mu$

$f(x) = \frac{1}{2}x^TAx + b^Tx + C$

$f^*(x) = \frac{1}{2}(y-b)^TA^{-1}(y-b) - C$

$= y^TA^{-1}y \quad b=0\ C=0$

$\max_\mu -\frac{1}{4}\mu^TAW^{-1}A^T\mu - b^T\mu$

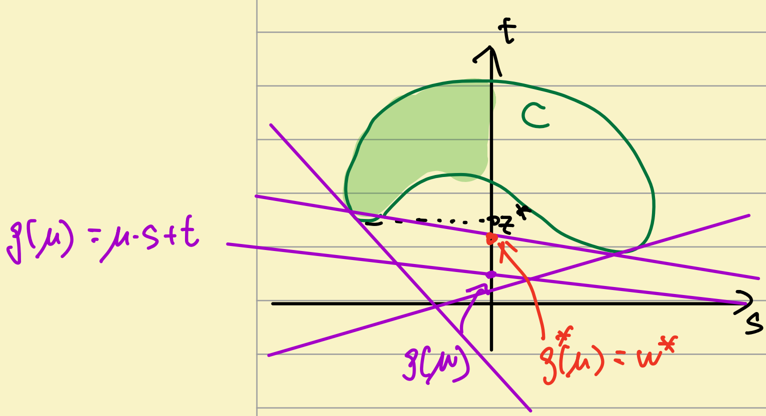
s.t. $\mu \ge 0$

对偶问题的几何解释:

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

$C \subset \mathbb{R}^n$ 是一个子集 $\quad C=\{(g(x), f(x)) \mid x\in\mathbb{R}^n\}$

最优解 $\quad z^* = \inf\{t \mid \underbrace{(s,t)\in C,\ s\le 0}_{\text{可行集}}\}$

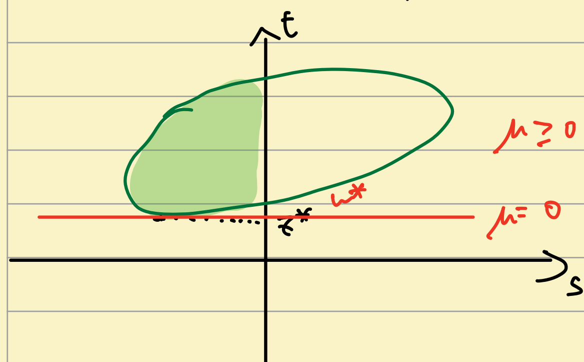


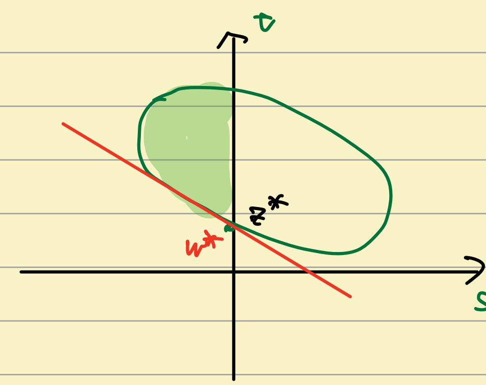
对偶问题:

$$g(\mu) = \inf_x \{f(x)+\mu g(x)\}$$
$$= \inf_{(s,t)\in C}\{\mu\cdot s+t\}$$

$$w^* = \sup_{\mu\ge 0} g(\mu)$$
$$= \sup_{\mu\ge 0}\ \inf_{(s,t)\in C}\{\mu\cdot s+t\}$$

凸问题: 强对偶条件成立





无约束问题的最优性理论

① 解的存在性.

② 如何设计算法求最优解.

$$\min f(x) \quad s.t.\ x \in \Omega \qquad (1)$$

$f(x)=\sqrt{x} \quad x\in(0,1] \quad \min f(x)$

Thm (Weierstrass Theorem) $f:\Omega \to (-\infty,+\infty]$ 适当且闭的函数. 若:

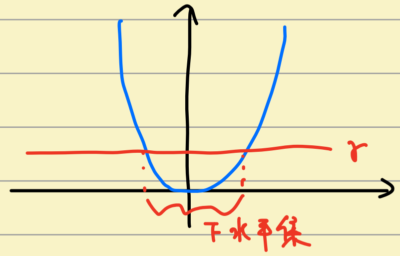
① $\mathrm{dom} f = \{x\in\Omega, f(x)<+\infty\}$ 是有界的

② 存在一个 r 下水平集: $C_r=\{x\in\Omega, f(x)\le r\}$ 是非空且有界的.

③ f 是强制的 (coersive): $\forall \|x^k\|\to+\infty$, $x^k\in\Omega$, 有
$$\lim_{k\to\infty} f(x^k)=+\infty$$

三个条件其中一个满足, 那么优化问题 (1) 的最小值点集
$$\{x\in\Omega \mid f(x)\le f(y), \forall y\in\Omega\}$$
是非空且紧的.

E.g. $f(x)=x^2 \quad x\to+\infty \quad f(x)\to+\infty$



一阶最优性条件

Defn (下降方向) 对于函数 f 是可微的, 如果存在向量 d 满足
$$\nabla f(x)^T d<0$$
那么称 d 为 f 在 x 点处的一个下降方向.

$$x^{k+1}=x^k+\alpha d$$

$$f(x+\alpha d)<f(x)$$

如果 f 在 x 点处存在一个下降方向 d，那么对于任意 $T>0$，存在 $t\in(0,T]$
使得

$$f(x+td) < f(x)$$

Proof：考虑 $g(t) = f(x+td)$

下降方向

$$g'(t) = \nabla f(x+td)^T d$$

$$g'(0) = \lim_{t\to 0}\frac{g(t)-g(0)}{t} = \lim_{t\to 0}\frac{f(x+td)-f(x)}{t} = \nabla f(x)^T d < 0$$

$$\Rightarrow f(x+td) - f(x) < 0 \Rightarrow f(x+td) < f(x)$$