约束优化问题（原问题）

$$\begin{cases} \min\limits_{x \in \mathbb{R}^P} f(x) \quad \text{Primal problem} \\ s.t. \quad m_i(x) \le 0, \quad i = 1, \cdots, M \\ \qquad n_j(x) = 0, \quad j = 1, \cdots, N \end{cases}$$

等价

写成 Lagrange 乘子法函数形式：

$$\mathcal{L}(x, \lambda, \eta) = f(x) + \sum_{i=1}^{M} \lambda_i m_i(x) + \sum_{j=1}^{N} \eta_j n_j(x)$$

$$\begin{cases} \min\limits_{x} \max\limits_{\lambda, \eta} \mathcal{L}(x, \lambda, \eta) \\ s.t. \quad \lambda_i \ge 0 \end{cases} \Rightarrow x \in \left\{ \underset{x}{\arg} \{ m_i(x) \le 0 \} \right\}$$

{好的 x 的集合}

如果 x 违反了约束 $m_i(x)$，则 $m_i(x) \ge 0$，

$\max\limits_{\lambda} \mathcal{L} \to \infty$；如果 x 符合约束，$\max\limits_{\lambda} \mathcal{L} \ne +\infty$

$$\min_{x} \max_{\lambda} \mathcal{L} = \min_{x} \{ \underbrace{\max \mathcal{L}}_{\arg m_i(x) \le 0}, +\infty \}$$

（原问题是关于 x 的函数）

对偶问题

$$\begin{cases} \max\limits_{\lambda, \eta} \min\limits_{x} \mathcal{L}(x, \lambda, \eta) \\ s.t. \quad \lambda_i \ge 0 \end{cases}$$

（对偶问题是关于 λ, η 的函数）

弱对偶性： ch6 重点 ☆☆☆

对偶问题 ≤ 原问题

$$d \le p$$

proof: max min ≤ min max

(⇒)

$$\max_{\lambda, \eta} \min_{x} \mathcal{L}(x, \lambda, \eta) \le \min_{x} \max_{\lambda, \eta} \mathcal{L}(x, \lambda, \eta)$$

proof:

$$\underbrace{\min_{x} \mathcal{L}(x, \lambda, \eta)}_{A(\lambda, \eta)} \le \mathcal{L}(x, \lambda, \eta) \le \underbrace{\max_{\lambda, \eta} \mathcal{L}(x, \lambda, \eta)}_{B(x)}$$

□

ch5 notes

对偶性的几何解释★

$$\begin{cases} \min\limits_{x\in R^n} f(x) \\ s.t.\ m_1(x) \le 0 \end{cases}$$

D: domain 定义域

$= \mathrm{dom} f \cap \mathrm{dom}\, m_1$

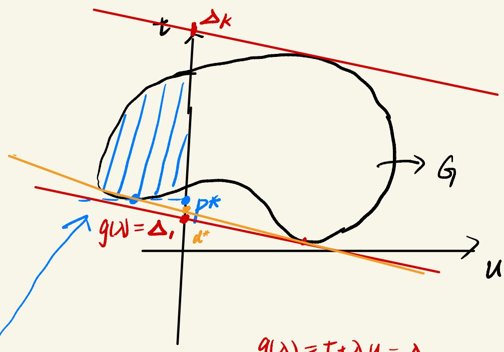
$L(x,\lambda) = f(x) + \lambda m_1(x),\ \lambda \ge 0$

$p^* = \min f(x)$ (原问题最优解)

$d^* = \max\limits_{\lambda} \min\limits_{x} L(x,\lambda)$ (对偶最优解)

$G = \{(m_1(x), f(x)) \mid x\in D\}$

$= \{(u,t) \mid x \in D\}$



弱对偶

$g(\lambda) = t + \lambda u = \Delta_1$

两点相切使g(λ)最大（调整λ）

$p^* = \inf\{t \mid (u,t)\in G,\ u \le 0\}$

$d^* = \max\limits_{\lambda} \min\limits_{x} L(x,\lambda)$

$= \max\limits_{\lambda} \underbrace{\min\limits_{x}(t+\lambda u)}_{g(\lambda)}$

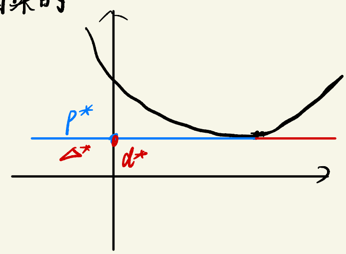
$= \max\limits_{\lambda} g(\lambda)$

$\Downarrow$

$g(\lambda) = \min(t+\lambda u)$

$= \inf\{t+\lambda u \mid (u,t)\in G\}$

而凸集时



强对偶

凸优化 + slater条件（充分非必要条件）

$\Rightarrow d^* = p^*$

SVM：二次规划问题（天生满足slater条件）

Slater Condition（Slater条件） Ch7

$\exists\ \tilde{x} \in \text{relint } D,\ s.t.\ \forall i=1,\cdots,M,\ m_i(\tilde{x}) < 0$

relint: relative interior（相对内部）

① 对于大多数凸优化，slater条件成立

② 放松的slater：M中有k个仿射函数，

M−k个非仿射函数

Convex + Slater $\Rightarrow$ Strong Duality

$\Updownarrow$

KKT

$p^* \to x^*$

$d^* \to \lambda^*, \eta^*$

Ch8

KKT Condition

KKT:

- 可行条件
  - $m_i(x^*) \le 0$
  - $n_j(x^*) = 0$ → primal feasibility
  - $\lambda^* \geq 0$ → dual feasibility
- 互补松弛：$\lambda_i^* m_i = 0$ → complementary slackness
- 梯度为0：$\left.\dfrac{\partial \mathcal{L}(x,\lambda^*,\eta^*)}{\partial x}\right|_{x=x^*} = 0$ → stationarity of the Lagrangian with respect to x

$$d^* = \max_{\lambda,\eta} g(\lambda,\eta) = g(\lambda^*,\eta^*)$$

$$= \min_x \mathcal{L}(x,\eta^*,\lambda^*)$$

① $$\le \mathcal{L}(x,\lambda^*,\eta^*)$$

$$= f(x^*) + \sum_i \underbrace{\lambda_i^*}_{\ge 0} \underbrace{m_i}_{\le 0} + \sum_j \eta_j^* \underbrace{\cancel{n_j}}_{=0}$$

② $$\le f(x^*)$$

$$= p^*$$

②

$$\Rightarrow \sum_i \lambda_i^* m_i = 0$$

$$\Rightarrow \lambda_i m_i = 0,\ \forall i=1,2,\cdots,M$$

①

$$\min \mathcal{L}(x^*,\lambda^*,\eta^*)$$

$$= \mathcal{L}(x^*,\lambda^*,\eta^*)$$

$\mathcal{L}(x,\lambda^*,\eta^*) = \underline{\min \mathcal{L}}$ 极小值